

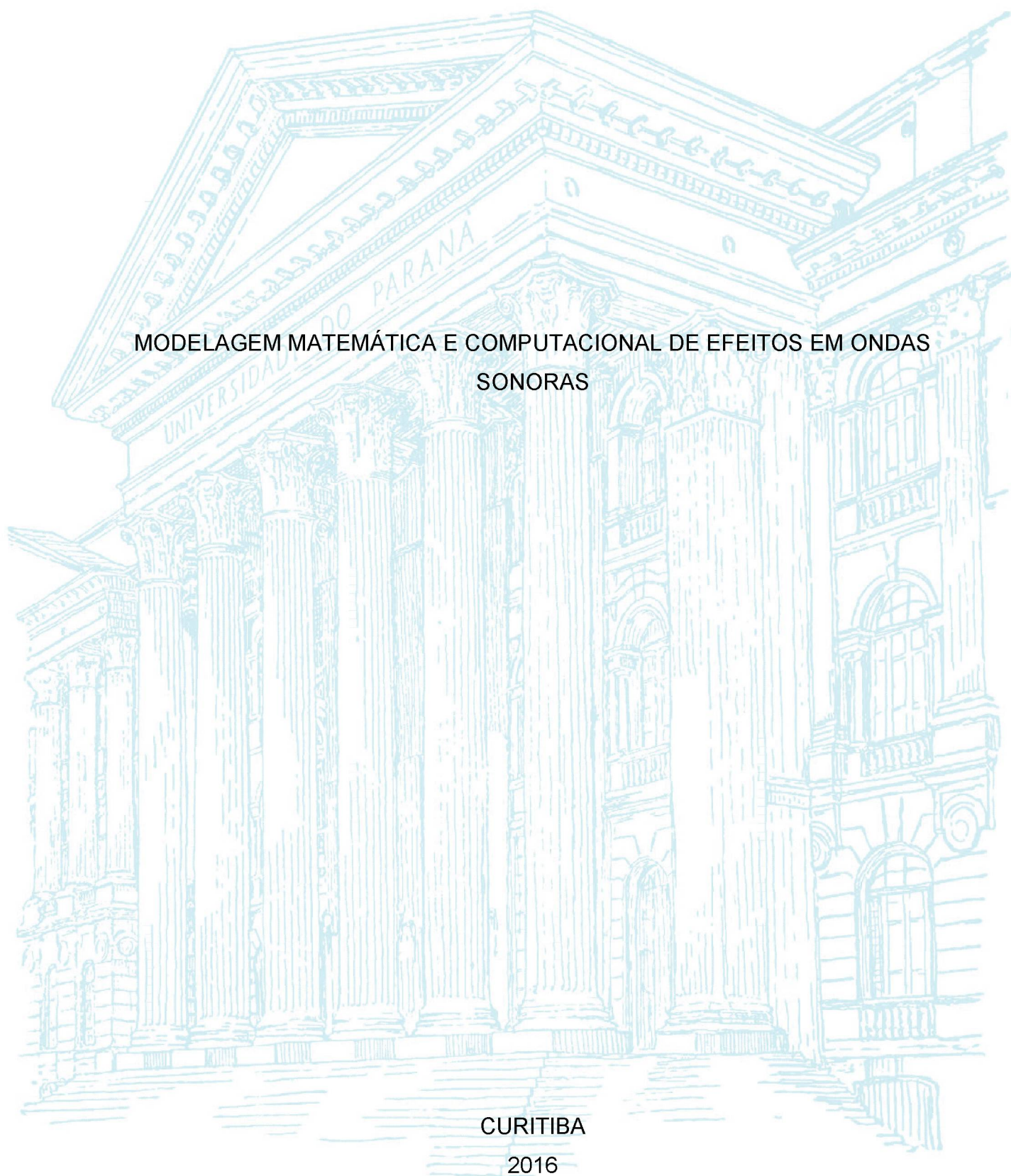
UNIVERSIDADE FEDERAL DO PARANÁ

CARLOS CESAR DE CARVALHO COVO

MODELAGEM MATEMÁTICA E COMPUTACIONAL DE EFEITOS EM ONDAS  
SONORAS

CURITIBA

2016



CARLOS CESAR DE CARVALHO COVO

MODELAGEM MATEMÁTICA E COMPUTACIONAL DE EFEITOS EM ONDAS  
SONORAS

Dissertação apresentada ao curso de Mestrado Profissional em Matemática em Rede Nacional (PROFMAT), Setor de Ciências Exatas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Matemática.

Orientadora: Prof.<sup>a</sup> Dra. Paula Rogeria Lima Couto

CURITIBA

2016



Catálogo na Fonte: Sistema de Bibliotecas, UFPR  
Biblioteca de Ciência e Tecnologia

C873m

Covo, Carlos Cesar de Carvalho  
Modelagem matemática e computacional de efeitos em ondas sonoras  
[recurso eletrônico] / Carlos Cesar de Carvalho Covo. – Curitiba, 2016.

Dissertação - Universidade Federal do Paraná, Setor de Ciências Exatas,  
Programa de Pós- Graduação em Matemática em Rede  
Nacional(PROFMAT), 2016.

Orientadora: Paula Rogeria Lima Couto.

1. Ondas sonoras. 2. Computação – Matemática. 3. Scilab (Programa de  
computador). 4. Modelos matemáticos. I. Universidade Federal do Paraná. II.  
Couto, Paula Rogeria Lima. III. Título.

CDD: 511.8

Bibliotecária: Vanusa Maciel CRB- 9/1928



MINISTERIO DA EDUCAÇÃO  
SETOR DE CIÊNCIAS EXATAS  
UNIVERSIDADE FEDERAL DO PARANÁ  
PRO-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO MATEMÁTICA EM REDE  
NACIONAL - 31075010001P2

## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em MATEMÁTICA EM REDE NACIONAL da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **CARLOS CESAR DE CARVALHO COVO** intitulada: **MODELAGEM MATEMÁTICA E COMPUTACIONAL DE EFEITOS EM ONDAS SONORAS**, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua aprovação no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 26 de Fevereiro de 2016.

PAULA ROGERIA LIMA COUTO

Presidente da Banca Examinadora (UNIVERSIDADE FEDERAL DO PARANÁ)

ALEXANDRE LUIS TROVON DE CARVALHO

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

ALISSON ANTONIO MARTINS

Avaliador Externo (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ)

## **AGRADECIMENTOS**

Agradeço a minha família, em especial a minha esposa e a minha filha, pelo apoio que tive durante o curso. Agradeço aos professores do curso e a instituição de ensino UFPR, pelo conhecimento ofertado. Agradeço também a CAPES pela bolsa de estudo concedida e a SEED por conceder o afastamento com ônus das minhas atividades laborais, sem isso seria inviável a realização deste trabalho.



## RESUMO

O presente trabalho tem como objetivo modelar matematicamente e computacionalmente alguns efeitos sonoros a partir de um som gravado digitalmente, quantificando-o em valores numéricos e transformando esses valores através de funções matemáticas que alterarão o som original. Dentre os efeitos que serão modelados têm-se: a variação de volume sonoro; o envelopamento dinâmico; delays, reverberações e ecos; variação da frequência de um som harmônico associado a uma nota musical; reconstrução sintetizada de um sinal e composição de sons harmônicos. Utiliza-se para tanto implementações computacionais de modelos temporais a partir do processamento digital do sinal sonoro. Este trabalho tem também como objetivo compreender o fenômeno sonoro, relacionando esse fenômeno com a percepção que temos do mesmo, como ele é gerado, como se propaga e como é captado. Estuda-se para isso os diversos parâmetros que permeiam esse fenômeno e como ocorre o processamento digital do sinal sonoro. Como resultado são apresentados vários algoritmos desenvolvidos no Scilab, que exploram uma matemática básica e que representam na prática um sintetizador de áudio. Palavras-chave: Modelagem matemática, ondas sonoras, sintetizador de áudio. .

Palavras-chave: Modelagem matemática, ondas sonoras, sintetizador de áudio.

## **ABSTRACT**

The following work aims to model mathematically and computationally some effects sounds from a digital recorded sound, quantifying it in numerical values and transforming these through mathematical functions that will change the original sound. Among the effects that will be modeled, there is: a sound volume variation; dynamic envelopment; delays, reverberations and echoes; frequency variation an harmonic sound associated with a musical note; synthesized reconstruction of a signal and composition of harmonic sounds. It is used for this computational implementations of temporal models from processing signal sound. This work also aims to understand the phenomenon of sound, relating this phenomenon to the perception that we have for it, how it is generated, how it propagates and how it is captured. For this study, there is a parameter that permeates this phenomenon, and how the digital processing signal occurs. As results, several algorithms developed are showed in the Scilab, which explore a basic math and which represent audio synthesis.

**Keywords:** Mathematical modeling, sound waves, audio synthesizer

# Sumário

<b>1 INTRODUÇÃO.....</b>	<b>1</b>
<b>2 FUNDAMENTAÇÃO TEÓRICA.....</b>	<b>3</b>
2.1 ONDA SONORA.....	4
2.1.1 Formato de propagação.....	5
2.1.2 Direção de vibração e de propagação de ondas.....	5
2.1.3 Velocidade do som.....	6
2.1.4 Comprimento de onda longitudinal ou de compressão.....	8
2.1.5 Período.....	9
2.1.6 Frequência.....	9
2.1.7 Ondas sonoras em um movimento harmônico simples.....	9
2.1.8 Amplitude.....	10
2.1.9 Intensidade do som.....	11
2.1.10 Interferência de ondas.....	12
2.1.11 Interferência de ondas.....	12
2.2 PERCEPÇÃO DO SOM (PSICOACÚSTICA).....	14
2.2.1 A audição humana.....	15
2.2.2 Relação qualitativa entre sensação e estímulo.....	15
2.2.3 Nível de intensidade sonora.....	16
2.2.4 Espectro sonoro.....	19
2.2.5 Altura.....	19
2.2.6 Curvas isofônicas.....	20
2.3 ANÁLISE ESPECTRAL.....	21
2.4 PROCESSAMENTO DE SINAIS DE ÁUDIO .....	23
<b>3 MODELAGEM MATEMÁTICA E COMPUTACIONAL DE EFEITOS EM ONDAS SONORAS.....</b>	<b>30</b>
3.1 REPRESENTAÇÃO MATEMÁTICA DOS DADOS INICIAIS.....	31
3.2 VARIAÇÃO DO VOLUME DO SINAL DE ÁUDIO.....	32
3.2.1 Algoritmo no Scilab para a variação do volume de áudio no formato .wav.....	32
3.2.2 Aplicação do algoritmo para a variação do volume de áudio.....	33
3.3 ENVELOPAMENTO DINÂMICO.....	34
3.3.1 Envelopamento do sinal de áudio por função afim.....	35



3.3.1.1 Algoritmo no Scilab para envelopamento por função afim de um áudio no formato .wav.....	39
3.3.1.2 Resultado da aplicação do algoritmo para envelopamento de um áudio por função afim.....	40
3.3.2 Envelopamento do sinal de áudio por função exponencial.....	41
3.3.2.1 Algoritmo no Scilab para envelopamento por função exponencial de um áudio no formato .wav.....	42
3.3.2.2 Resultado da aplicação do algoritmo para envelopamento de um áudio por função exponencial.....	42
3.3.3 Envelopamento do sinal por funções mais flexíveis.....	43
3.3.3.1 Implementação no Scilab para a captura de envelope dinâmico.....	44
3.3.3.2 Resultado da aplicação do algoritmo para a captura de envelope dinâmico.....	46
3.4 EFEITOS DE ÁUDIO PARA REVERBERAÇÃO, DELAY E ECO.....	47
3.4.1 Implementação no Scilab para modelagem do efeito reverb.....	51
3.4.2 Descrição do algoritmo de modelagem do efeito reverb.....	52
3.4.3 Aplicação do algoritmo para modelagem do efeito reverb.....	53
3.4.4 Considerações sobre o modelo matemático do efeito reverb .....	54
3.5 MODELAGEM DE ONDAS SONORAS APLICADA À MÚSICA.....	54
3.5.1 Sintetização de uma escala musical temperada.....	56
3.5.2 Implementação no Scilab para sintetização de nota musical com envelope exponencial .....	57
3.5.3 Descrição do algoritmo da implementação no Scilab para sintetização de nota musical com envelope exponencial.....	58
3.5.4 Aplicação da implementação no Scilab para sintetização de nota musical com envelopamento exponencial.....	58
3.6 SAÍDA DE ÁUDIO.....	60
<b>REFERÊNCIAS .....</b>	<b>72</b>
<b>ANEXO A – ROTEIRO PARA A 1ª PARTE DO PRÉ-PROCESSAMENTO:</b>	
<b>CAPTURA DO SINAL DE ÁUDIO.....</b>	<b>63</b>
<b>ANEXO B – ROTEIRO PARA A 2ª PARTE DO PRÉ-PROCESSAMENTO:</b>	
<b>IMPORTAÇÃO DOS DADOS DE ÁUDIO NO SCILAB.....</b>	<b>67</b>
<b>ANEXO C – ROTEIRO PARA A GERAÇÃO DE UM ARQUIVO DE ÁUDIO NO SCILAB.....</b>	<b>72</b>

# Modelagem Matemática e Computacional de Efeitos em Ondas Sonoras

por

Carlos Cesar de Carvalho Covo

**Orientadora do Trabalho:**

Prof.<sup>a</sup> Dra. Paula Rogéria Lima Couto

Disponível via INTERNET:  
<http://www.mat.ufpr.br>

# Modelagem Matemática e Computacional de Efeitos em Ondas Sonoras

*Carlos Cesar de Carvalho Covo*

Departamento de Matemática - UFPR  
81531-980, Curitiba, PR  
Brasil

e-mail: carloscovo@ufpr.br

30 de Março de 2016

## Resumo

O presente trabalho tem como objetivo modelar matematicamente e computacionalmente alguns efeitos sonoros a partir de um som gravado digitalmente, quantificando-o em valores numéricos e transformando esses valores através de funções matemáticas que alterarão o som original. Dentre os efeitos que serão modelados têm-se: a variação de volume sonoro; o envelopamento dinâmico; delays, reverberações e ecos; variação da frequência de um som harmônico associado a uma nota musical; reconstrução sintetizada de um sinal e composição de sons harmônicos. Utiliza-se para tanto implementações computacionais de modelos temporais a partir do processamento digital do sinal sonoro. Este trabalho tem também como objetivo compreender o fenômeno sonoro, relacionando esse fenômeno com a percepção que temos do mesmo, como ele é gerado, como se propaga e como é captado. Estuda-se para isso os diversos parâmetros que permeiam esse fenômeno e como ocorre o processamento digital do sinal sonoro. Como resultado são apresentados vários algoritmos desenvolvidos no Scilab, que exploram uma matemática básica e que representam na prática um sintetizador de áudio.

**Palavras-chave:** Modelagem matemática, ondas sonoras, sintetizador de áudio.

## 1 Introdução

Os sons em geral desempenham um papel importante para o desenvolvimento da espécie humana e da sociedade. Sabemos da existência dos sons graças a nossa audição que combinada com outros sentidos nos fornece uma enorme fonte de



informações, estabelecendo uma ponte entre o mundo e a maneira que interagimos com ele. O som é capaz de proporcionar e alterar emoções, através do som é que temos a música, que é capaz de alimentar a imaginação, prender a atenção e favorecer de forma eficiente o processo de ensino-aprendizagem. Desde criança ouvimos e interagimos com o som, é algo natural e instintivo, embora passamos boa parte das nossas vidas ouvindo e interagindo com uma imensa diversidade sonora, poucas vezes paramos para pensar a respeito. O que é o som? Como se forma? Como é possível haver tanta variedade sonora? Por que alguns sons soam agradáveis e outros não? Afinal, o que é o som?

Desde tempos remotos, do surgimento da linguagem oral, o ser humano vem aprimorando o uso do som para o seu desenvolvimento. Na atualidade podemos facilmente ouvir, compartilhar e modificar diversas músicas e sons gerados por instrumentos reais ou virtuais. Fazemos isso de forma quase automática com o auxílio de um computador ou de um smartfone, sem pensarmos, na maioria das vezes, como toda essa magia pode acontecer, no que está sustentado esta tecnologia. Presenciamos hoje uma era digital, difícil para muitos acompanharem, mas que é desfrutada por quase todos. A nossa televisão não recebe mais um sinal analógico, temos agora um sinal digital, a qualidade superior é visível, novas tecnologias surgem para substituir outras antes mesmo que possamos compreendê-las. Grande parte dos sons que ouvimos não é mais natural, são sons sintetizados eletronicamente, presentes em alarmes de celulares, beeps de micro-ondas, teclados musicais e em diversos outros aparelhos eletrônicos. Os sons gerados por esses aparelhos escondem um longo processo de desenvolvimento científico, não surgiram de repente, teve a participação de matemáticos, físicos, engenheiros, que cada um em seu tempo, contribuiu para que esse desenvolvimento fosse possível. Nesse processo toda a matemática sempre esteve presente, sem ela nada disso seria possível, ela é a linguagem, a teoria, a forma pela qual é possível acompanhar todo esse progresso científico.

A tecnologia da qual dispomos nos dias de hoje é sustentada por modelos físicos e matemáticos que descrevem com precisão os diversos fenômenos da natureza, em particular, os fenômenos sonoros. Fenômenos como eco, reverberação ou efeitos criados, como sons modificados com sintetizadores podem ser compreendidos a partir de modelos matemáticos que usam uma matemática de nível médio, que envolvem por exemplo manipulação de matrizes, sequências numéricas, funções trigonométricas, funções reais e complexas, noções de limite e continuidade, amostragem, entre outros conceitos matemáticos que serão abordados nos modelos matemáticos elaborados neste trabalho.

O estudo do som, mais precisamente das ondas sonoras, se apresenta como um importante e interessante argumento para alavancar e conectar diversos conceitos matemáticos de nível médio, relativamente simples de serem compreendidos e que podem abrir um leque de possibilidades para o professor, propiciando uma interdisciplinaridade com a física no trato do conteúdo de ondas e acústica.

Neste trabalho percorre-se um caminho que faz uma ponte entre o conheci-

mento teórico e o prático. Inicia-se com o estudo do conceito físico do som e da percepção que temos do mesmo. Apresenta-se algumas relações matemáticas que descrevem esses fenômenos, passando pelo processamento digital do sinal sonoro, que fornecerá os dados numéricos necessários para a elaboração dos modelos matemáticos que reproduzirão, de forma aproximada, alguns fenômenos e efeitos sonoros presentes no cotidiano, através de uma matemática relativamente simples.

O que motivou a realização deste trabalho são os resultados práticos obtidos pelos modelos computacionais. Esses modelos foram construídos de forma heurística, são relativamente simples e fáceis de entender, e facilitam a compreensão dos conceitos matemáticos que serão apresentados. Esses modelos consistem basicamente em uma representação e manipulação das ondas sonoras por matrizes em um ambiente computacional, utilizando o processamento de sinais de áudio.

Usa-se como ferramentas computacionais o software livre de edição e gravação de áudio Audacity<sup>1</sup>, o software livre científico de computação numérica Scilab<sup>2</sup> e o software livre de geometria dinâmica GeoGebra<sup>3</sup>, para a captação, edição e modelagem computacional dos sons gravados. Para facilitar e automatizar a análise dos dados fornecidos pelo processamento de sinais de ondas sonoras, e devido às limitações dos referidos softwares, utiliza-se na aplicação dos modelos computacionais ondas sonoras com extensão total curta, desde as mais simples, de comportamento previsível, como as definidas por senoides, até as mais aleatórias, como as geradas de forma natural.

## 2 Fundamentação teórica

Para entender o que é o som, apresentam-se a seguir, de forma geral, vários componentes que tratam da descrição da sua natureza, e os mesmos serão detalhados posteriormente nas seções indicadas entre parênteses.

A sensação sonora é uma qualidade perceptiva de vibrações das partículas em um meio material, essas vibrações são consequentes de sucessivas frentes de compressão mecânica nesse mesmo meio em um pequeno intervalo de tempo. Para o fenômeno sonoro ocorrer é necessário a presença de três elementos: emissor (fonte sonora), meio de propagação (material) e receptor (percepção do som).

O emissor sonoro é o responsável pela perturbação do meio, é o que provoca as variações de pressão iniciais. As compressões resultantes podem se apresentar de

---

<sup>1</sup>O Audacity é um software livre de gravação, edição e reprodução de áudio digital, possui diversos filtros para efeitos e ferramentas de análise gráfica, tem a opção de definir e exportar dados numéricos de um áudio, e é bastante dinâmico e intuitiva a sua utilização.

<sup>2</sup>O Scilab é um software científico de computação numérica com várias ferramentas desenvolvidas disponíveis, de fácil implementação, para construções de algoritmos dos mais simples aos mais complexos, com grande aplicabilidade educacional e em pesquisa científica.

<sup>3</sup>O GeoGebra é um software livre de geometria dinâmica, de fácil utilização e bastante difundido no meio educacional.

forma periódica ou não, podem ter um período único (Seção 2.1.5) ou variável, e podem ter amplitude máxima única em cada período ou variar ao longo do tempo (Seção 2.1.8), dependendo da uniformidade ou não das perturbações iniciais do meio.

Para que ocorram as ondas sonoras (Seção 2.1) o meio material perturbado pela fonte sonora precisa ser elástico, ou seja, a substância do meio deve ser capaz de transmitir a energia da onda. Sendo o receptor o ouvido humano, para perceber o som, as ondas sonoras devem ser definidas dentro de uma faixa de frequências audível (Seção 2.2.4), e deve ter um nível de amplitude sonora audível e suportável (Seção 2.2.3). A forma como ocorrem as sucessivas variações de pressão determina algumas características do som, que pode ser senoidal, composto, complexo ou ruído.

A mais simples das ondas sonoras é conhecida como onda senoidal, gerada por um “Movimento Harmônico Simples” (Seção 2.1.7). O som gerado por uma onda senoidal é um som obtido por uma única frequência (Seção 2.1.6) com amplitude máxima se repetindo periodicamente, e que pode ser descrito matematicamente por uma função senoidal do tempo. Porém, esse tipo de onda sonora não é natural, por ser improvável que exista uma fonte sonora com uma vibração com intensidade estritamente constante em uma frequência audível, e também por existir entre o emissor sonoro e o receptor um meio material não uniforme, devido ao movimento das partículas e variação da densidade, que, por menor que seja, resulta na variação do período e da amplitude máxima em distintos intervalos de tempo. Mesmo não sendo natural, o movimento harmônico simples pode ser, por exemplo, sintetizado eletronicamente no emissor sonoro.

Para todos os sons audíveis existem fontes sonoras que produzem vibrações periódicas, não estritamente periódicas e aperiódicas, e existem também fontes sonoras com intensidades de vibração constantes e não constantes. Ondas sonoras que não são totalmente estáveis em período e amplitude geram sons denominados tônicos ou compostos. De acordo com Flo Menezes em [9], os sons não senoidais são resultantes de sobreposições de sons senoidais em seu espectro sonoro. Um som é dito composto quando as senoides são parciais de seu espectro sonoro. Os sons harmônicos são sons compostos de senoides em proporção de números inteiros a alguma frequência fundamental. Um som complexo é um som composto não-harmônico, pode ser gerado por componentes senoidais em proporção de números não inteiros. O ruído é um fenômeno sonoro resultante de uma vibração aperiódica, onde não se tem um espectro discreto de frequências, o ruído não possui um modelo periódico definido [9].

## 2.1 Onda sonora

Uma onda sonora é uma onda mecânica que se propaga em um meio material elástico (sólido, líquido ou gasoso), onde ocorre a transferência de energia sem ocorrer a transferência de matéria, isto é, o que se propaga não são as partículas



do meio e sim o movimento [1], as partículas apenas oscilam próximas às suas posições de repouso. Uma onda sonora pode ser progressiva, quando se propaga de um lugar para outro, ou estacionária, quando se forma pela interferência de ondas que se propagam em sentidos opostos, condição típica de alguns instrumentos musicais de sopro.

A onda sonora é o resultado de um movimento oscilatório transmitido de partícula para partícula até chegar aos nossos ouvidos. O som, ou sensação sonora, não é o resultado de uma, mas de sucessivas ondas sonoras geradas por algum corpo em vibração. Esse fenômeno pode ser explicado pelo princípio de Huygens-Fresnel<sup>4</sup>, quando aplicado a ondas sonoras. A partir da vibração de uma fonte sonora agindo sobre uma partícula do meio material, e a mesma, ao vibrar, transmite para uma partícula adjacente a sua oscilação, se comportando como uma nova fonte sonora [10]. Entende-se como fonte sonora qualquer corpo que vibra perturbando o meio material em que se encontra, emitindo uma sucessão de ondas sonoras com alguma frequência e amplitude detectável ao ouvido humano [1].

### **2.1.1 Formato de propagação**

Quanto ao formato de propagação, as ondas podem se apresentar das seguintes formas: unidimensionais, quando se propagam em uma única direção, como é o caso de uma onda numa corda provocada por um movimento inicial transversal; bidimensionais, que são as ondas que se propagam em todas as direções de uma superfície, como as ondas circuncêntricas formadas por uma pedra jogada em um espelho d'água; e tridimensionais, que são ondas que se propagam em todas as direções, como as ondas sonoras.

Quando uma fonte sonora se encontra a uma distância suficientemente grande em comparação às suas próprias dimensões, ela pode ser considerada uma fonte puntiforme e, nessas condições, as frentes que partem dessa fonte se propagam como esferas concêntricas que se expandem em todas as direções a partir da fonte de perturbação. Portanto, pode-se supor que as ondas sonoras tem formatos esféricos desde que sua fonte seja considerada pontual e isotrópica [2].

### **2.1.2 Direção de vibração e de propagação de ondas**

Pode-se classificar a direção de vibração e de propagação de uma onda qualquer como: onda longitudinal, transversal ou mista. As ondas sonoras nos fluidos (meio material líquido ou gasoso) são longitudinais e possuem uma mesma direção de vibração e de propagação, isto é, a vibração da fonte geradora do som ocorre na mesma direção do movimento oscilatório. Conforme descrição em [3], nos sólidos,

---

<sup>4</sup>Princípio de Huygens - Christian Huygens (1629-1695): todo ponto de uma frente de onda se comporta como uma nova fonte de ondas, que se propagam, com a mesma frequência, para além da região já atingida pela onda original [11, 30].

as ondas mecânicas podem ser longitudinais ou transversais, mas nos fluidos (que não suportam forças de cisalhamento), só podem ser longitudinais.

A propagação gerada por uma vibração transversal ocorre quando a vibração da fonte geradora é perpendicular à direção de propagação da onda. Ondas transversais são ondas associadas a tensões tangenciais (cisalhamento) como é o caso das ondas eletromagnéticas.

Existe um outro tipo de propagação de onda, as provocadas na superfície da água por exemplo, que não são longitudinais e também não são transversais, essas ondas se propagam de forma circuncêntrica aproximadamente, com componentes perpendiculares e na direção da propagação, são também conhecidas como ondas bidimensionais ou superficiais, vibram longitudinalmente e transversalmente ao mesmo tempo, são as ondas mistas de direção de vibração. Tanto as ondas longitudinais como as transversais são classificadas como ondas progressivas [2].

### 2.1.3 Velocidade do som

O que gera o som são os corpos em vibração, e por se tratar de uma onda mecânica, a velocidade da propagação do som depende do meio material no qual se propaga e não do movimento inicial da fonte sonora. Em meios sólidos a velocidade de propagação é maior que em meios líquidos e maior ainda em um meio material gasoso (Tabela 1), ou seja, a densidade do meio material é que determina a velocidade do som. Como a temperatura influencia diretamente no volume, e a variação do volume influencia na densidade, veremos a seguir que a velocidade de propagação da onda sonora depende também da temperatura no meio material de propagação, sendo a variação mais expressiva em meios líquidos e gasosos do que em meios sólidos [6].

Meio material	Velocidade em $m/s$
Ar ( $0^{\circ}C$ )	332
Ar ( $20^{\circ}C$ )	344
Hélio	965
Hidrogênio	1284
Água ( $0^{\circ}C$ )	1402
Água ( $20^{\circ}C$ )	1482
Água do mar	1522
Alumínio	6420
Aço	5941
Granito	6000

Tabela 1: Velocidade de propagação do som. Fonte: adaptado de [4], p. 150.

Vamos nos restringir em analisar a velocidade do som  $\nu_s$  no meio gasoso, por ser o meio mais comum pelo qual percebemos o som. Quando uma fonte sonora

perturba um meio específico, cria-se uma frente de compressão que se desloca longitudinalmente percorrendo uma distância  $\Delta d$  durante um tempo  $\Delta t$ . Pode-se representar a velocidade do som nesse meio específico através da equação

$$\nu_s = \frac{\Delta d}{\Delta t},$$

isto é, a velocidade do som nesse meio é igual a razão entre a distância percorrida pela frente de compressão e o tempo utilizado para percorrê-la.

A especificidade do meio material pelo qual a onda sonora se propaga depende da densidade e da pressão do meio. Considerando o modelo de um gás ideal, isto é, concebendo um meio gasoso onde a temperatura está distante do ponto de liquefação e onde a pressão não seja alta, e considerando um sistema adiabático para esse meio material, a velocidade do som pode ser obtida da equação de onda no fluido<sup>5</sup>, dada por

$$\nu_s = \sqrt{\frac{\gamma \cdot P}{\rho}}, \quad (1)$$

onde  $P$  é a pressão no gás, que em condições normais de temperatura e pressão assume o valor  $P_0 = 1 \text{ atm} \approx 1,013 \times 10^5 \text{ N/m}^2$ , para a temperatura  $\bar{T} = 273 \text{ K}$  ( $0^\circ \text{C}$ );  $\rho$  é a densidade volumétrica ou massa específica em um ponto do espaço, dada pela razão entre a massa  $m$  do fluido, em  $\text{kg}$ , em um dado volume  $V$  do fluido, em  $\text{m}^3$ , isto é,  $\rho = \frac{m}{V}$ .

Segundo [6], a massa específica de um gás varia consideravelmente com a pressão, mas a massa específica de um líquido não varia. Portanto, apenas os gases são compressíveis. A massa específica do ar a  $20^\circ \text{C}$  e com  $1 \text{ atm}$  de pressão é de  $\rho = \rho_0 \approx 1,293 \text{ kg/m}^3$  [6]. O número  $\gamma$  representa a razão entre o calor específico  $C_p$  à pressão constante e o calor específico  $C_v$  ao volume constante,  $\gamma = \frac{C_p}{C_v}$ . Para gases ideais monoatômicos tem-se  $\gamma = 5/3 \approx 1,67$ , e para gases ideais diatômicos tem-se  $\gamma = 7/5 = 1,4$ , obtidos experimentalmente. Sendo o ar, basicamente, uma mistura de  $N_2$  e  $O_2$ , uma boa aproximação é  $\gamma \approx 1,4$  para o ar [2].

É possível relacionar a temperatura do meio material com a velocidade do som nesse meio. Partindo da equação de estado de um gás ideal

$$PV = nR\bar{T}, \quad (2)$$

que relaciona a pressão  $P$  no gás com o volume  $V$  e a temperatura  $\bar{T}$  em Kelvin, pode-se escrever  $P = \frac{nR\bar{T}}{V}$ , e sendo  $\rho = \frac{m}{V}$ , substituindo obtém-se

$$\frac{P}{\rho} = \frac{R\bar{T}}{\frac{m}{n}} = \frac{R\bar{T}}{M},$$

---

<sup>5</sup>Detalhes sobre a obtenção da equação de onda no fluido podem ser encontrados em [2].



onde  $M$  é a massa molar do gás dada pela razão entre a massa total do gás  $m$  e o número de *mols*  $n$ , isto é,  $M = \frac{m}{n}$  e  $R$  a constante de Rydberg dos gases ideais que vale  $8,314 \text{ J/mol} \cdot \text{K}$ . Assim, da Equação 1, para um gás ideal, tem-se

$$\nu_s = \sqrt{\frac{\gamma \cdot RT}{M}},$$

sendo  $R$  e  $M$  constante, pode-se definir  $\alpha = \sqrt{\frac{\gamma \cdot R}{M}}$  como sendo um valor específico constante característico do gás, portanto, sob as condições descritas acima, é possível relacionar a velocidade do som  $\nu_s$  em um determinado fluido com a temperatura  $\bar{T}$  do meio gasoso medida em Kelvin, por

$$\nu_s(\bar{T}) = \alpha \sqrt{\bar{T}}, \quad (3)$$

ou seja, a velocidade do som independe da pressão em um processo isotérmico em um gás ideal, mas varia com a raiz quadrada da temperatura [2].

Como um dos objetivos desse trabalho é sintetizar por modelagem comportamentos de ondas sonoras e também as sensações provocadas pelas mesmas no ouvinte em diversas situações onde naturalmente ocorre o fenômeno sonoro, é importante conhecer a velocidade do som em diversas condições que afetam o meio pelo qual as ondas se propagam. A velocidade do som é de fundamental importância para determinar, por exemplo, a acústica de um ambiente, para simular os fenômenos eco e reverberação, e também para determinar o atraso sonoro provocado pela diferença de tempo entre o som percebido por um ouvinte próximo da fonte sonora e de um outro ouvinte a uma maior distância determinada. O conhecimento da velocidade do som em um ambiente serve também para simular efeitos específicos, como por exemplo, a partir de um som monoaural, simular um som estereofônico, aplicando um atraso de tempo que simula a sutil diferença entre a chegada do som a um ouvido antes do outro, o que contribui para a sensação de profundidade e direcionamento do som<sup>6</sup> [2, 6, 9].

#### 2.1.4 Comprimento de onda longitudinal ou de compressão

Uma onda sonora, provocada por um movimento de perturbação no meio material, se propaga longitudinalmente através da variação de pressão, ou seja, a onda sonora se propaga na mesma direção do movimento perturbador do meio material. Após a ação da fonte sonora, tem-se como resultado sucessivas regiões alternadas de compressão e rarefação das partículas do meio material em relação a seu estado normal de pressão. O comprimento  $\lambda$  de uma onda sonora representa a menor distância entre dois pontos de compressão máxima ou rarefação máxima da pressão desse meio material (Figura 1).

---

<sup>6</sup>As demonstrações das equações presentes nesta seção e demais detalhes podem ser encontrados em [2], capítulo 6.

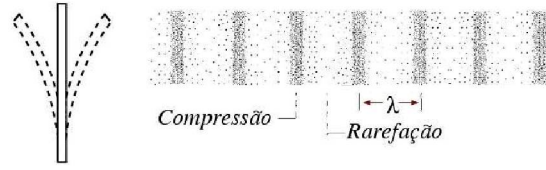


Figura 1: Representação de onda de compressão. Fonte: adaptado de [7].

### 2.1.5 Período

O período  $\tau$  é o tempo decorrido para que a pressão efetue o ciclo definido pelo comprimento de onda. Pode-se relacionar a velocidade do som  $\nu_s$ , com o comprimento  $\lambda$  e o período  $\tau$  por

$$\nu_s = \frac{\lambda}{\tau}. \quad (4)$$

### 2.1.6 Frequência

A frequência  $f$  de uma onda sonora indica o número de ciclos realizados por unidade de tempo, quando a unidade de tempo é o segundo, representamos a frequência em Hertz ( $Hz$ ). Desta forma, o período  $\tau$  é o inverso da frequência  $f$ , isto é,

$$\tau = \frac{1}{f}. \quad (5)$$

### 2.1.7 Ondas sonoras em um movimento harmônico simples

Para compreender o movimento oscilatório que descreve o som, é tradicionalmente recomendado analisar a princípio a forma mais simples de explicar o fenômeno vibratório, que é o movimento harmônico simples. Conforme consta em [6], todo movimento que se repete em intervalos regulares é chamado de movimento periódico ou movimento harmônico. O movimento harmônico simples é um movimento periódico que pode ser representado por uma função senoidal do tempo, o som gerado por um movimento harmônico simples é conhecido como tom puro. Desse modo, um tom puro ou tom senoidal é o som gerado por uma única frequência e pode ser obtido por uma função seno ou cosseno, onde os pontos da função com valor máximo representam as regiões de compressão máxima, e os mínimos os pontos de rarefação máxima [6, 9]. A Figura 2 ilustra a relação entre os máximos e mínimos de uma função senoidal com as regiões de compressão e rarefação máximas, respectivamente. Pode-se notar que o eixo das ordenadas no gráfico da função senoidal representa os valores da variação da pressão, enquanto o eixo das abscissas representa as coordenadas espaciais.



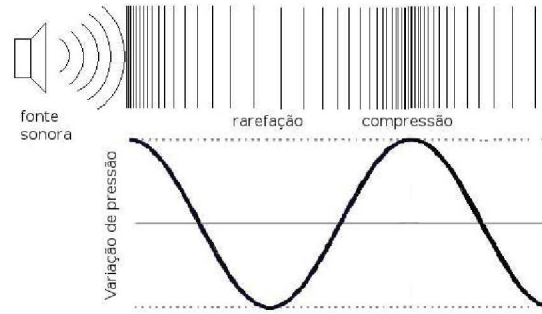


Figura 2: Representação de onda de compressão. Variação de pressão no espaço percorrido pela onda. Fonte: adaptado de [12].

### 2.1.8 Amplitude

A amplitude sonora representa a magnitude de oscilação de uma onda sonora. A amplitude máxima pode ser determinada pela diferença entre a pressão máxima (compressão máxima) e a pressão normal do meio material onde a onda se propaga, ou pela diferença da pressão normal e a pressão mínima (rarefação máxima). A amplitude sonora pode ser expressa por uma medida de pressão, dada em  $Pa$  ( $1 Pa = 1 N/m^2$ ). Considerando uma onda sonora senoidal, isto é, que se propaga em um movimento harmônico simples, a amplitude máxima ocorre em períodos regulares ao longo do tempo, e pode ser representada como o fator  $\Delta P_m$  da função

$$\Delta P(x, t) = \Delta P_m \sin(kx - \omega t), \quad (6)$$

cujo argumento da função seno representa o termo oscilatório.

A Equação 6 representa graficamente uma onda em movimento harmônico simples, onde  $\Delta P(x, t)$  é a variação da pressão em função da posição  $x$  do caminho percorrido pela onda e do instante  $t$ ;  $\Delta P_m$  é a constante que dá a máxima variação de pressão associada à onda;  $k$  é a constante que indica o número de comprimentos de onda por unidade de distância,  $k = \frac{2\pi}{\lambda}$ , onde  $\lambda$  é o comprimento de onda;  $\omega$  é a constante que representa a frequência angular,  $\omega = 2\pi f$ , onde  $f$  é a frequência;  $(kx - \omega t)$  é o termo oscilatório ou fase da onda. A Figura 3 mostra um gráfico que representa a Equação 6 para uma variação de pressão de uma onda sonora de  $1000 Hz$  com amplitude no limiar da dor.

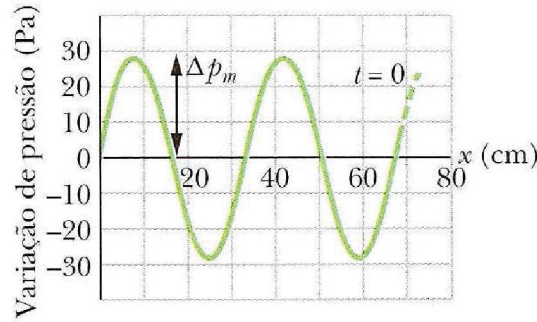


Figura 3: Gráfico de variação de pressão de uma onda sonora de  $1000\text{ Hz}$  com amplitude no limiar da dor. Fonte: adaptado de [6], p. 155.

A ocorrência periódica da amplitude máxima no espaço de propagação é algo difícil de acontecer naturalmente, devido a variação de densidade do meio material.

### 2.1.9 Intensidade do som

O que nos permite identificar o volume sonoro é a intensidade do som, que depende da transferência de energia da onda sonora. Considerando a onda sonora como uma onda plana, ou seja, uma onda que percorre um meio com densidade constante em cada plano e que tem sempre a mesma direção de propagação, a potência  $\bar{P}$ , que atravessa ou é absorvida por uma superfície de área  $S$ , aumenta proporcionalmente ao aumentar a área  $S$  da superfície, isto é,  $\bar{P} = I \cdot S$ . O fator  $I$  é definido como intensidade sonora, isto é, como a potência sonora absorvida por unidade de área de uma superfície. Sendo a intensidade sonora definida como potência por área, ela pode ser medida em watt por metro quadrado ( $\text{W/m}^2$ ). A intensidade sonora pode também ser expressa como o valor médio de energia por unidade de área perpendicular à direção de propagação.

Denomina-se pressão sonora  $p$  as pequenas variações de pressão no ar (ou outro meio) em relação a pressão atmosférica (ou pressão normal do meio), ela varia tanto no tempo quanto no espaço [5]. Considerando a componente da força que ocorre no mesmo sentido e direção de propagação da onda, e supondo que a intensidade não varia na superfície de área  $S$ , é possível escrever a intensidade como

$$I = \frac{\text{potência}}{\text{área}} = \frac{\text{energia}}{\text{área} \times \text{tempo}} = \underbrace{\frac{\text{força}}{\text{área}}}_p \times \underbrace{\frac{\text{distância}}{\text{tempo}}}_{v_s}, \quad (7)$$

isto é, a intensidade pode ser obtida pelo produto da pressão sonora pela velocidade de propagação da onda,  $I = p \cdot v_s$  [13].

A intensidade sonora varia naturalmente com a distância, supondo que a onda sonora se propaga de forma esférica e que a energia mecânica seja conservada

durante a propagação, da Equação 7 conclui-se que a intensidade sonora em um determinado ponto a uma distância  $r$  da fonte sonora com potência  $\bar{P}$  é dada por  $I = \frac{\bar{P}}{4\pi r^2}$ , sendo  $4\pi r^2$  a área da superfície de uma esfera de raio  $r$ . Desta forma, a intensidade sonora diminui com o quadrado da distância  $r$  da fonte sonora. Para medir a intensidade sonora  $I$  usam-se captadores de variação de pressão, ou seja, existe uma relação entre a intensidade sonora e a pressão sonora, considerando as características do meio de propagação ( $\rho$  e  $\nu_s$ ), esta relação é dada por

$$I = \frac{1}{2} \cdot \frac{\Delta P_m^2}{\rho \nu_s}, \quad (8)$$

onde,  $\Delta P_m$  é a amplitude máxima de uma onda sonora em movimento harmônico simples. Portanto, se duas ondas se propagarem em um mesmo meio, a que tiver a maior amplitude de pressão terá também a maior intensidade sonora. Porém, pode acontecer de duas ondas com mesma amplitude terem intensidades sonoras diferentes, é o caso de se propagarem em meios diferentes, ou seja, se possuírem distintas velocidades de propagação [2].

### 2.1.10 Interferência de ondas

O som é uma percepção consequente de objetos em vibração que não produzem apenas uma, mas várias ondas sonoras, que, em geral, não são geradas por um movimento harmônico simples, contudo, será admitido esse movimento para a análise da interferência. A interferência é o fenômeno que representa a superposição de ondas que se interagem em um ponto do espaço de propagação. Essa interferência acontece através da soma de cada onda senoidal que compõem cada uma das ondas sonoras. Quando duas ou mais ondas incidem simultaneamente em um único ponto (por exemplo, o ouvido humano), o deslocamento total nesse ponto é igual a soma algébrica dos deslocamentos das ondas individuais, esse fenômeno também pode ser chamado de superposição ou sobreposição de ondas.

A interferência de duas ondas

$$y_1(x, t) = A \sin(kx - \omega t)$$

e

$$y_2(x, t) = A \sin(kx - \omega t + \phi)$$

resulta na onda

$$y(x, t) = y_1(x, t) + y_2(x, t) = A \sin(kx - \omega t) + A \sin(kx - \omega t + \phi) \quad (9)$$

sendo  $\phi$  a diferença de fase entre as ondas. Através da identidade trigonométrica

$$\sin(a) + \sin(b) = 2 \cos\left(\frac{a-b}{2}\right) \sin\left(\frac{a+b}{2}\right)$$

obtém-se

$$y(x, t) = \underbrace{2A \cos\left(\frac{\phi}{2}\right)}_{\text{amplitude}} \underbrace{\sin\left(kx - \omega t + \frac{\phi}{2}\right)}_{\text{fase}}$$

onde: se  $\phi = 0$ , então  $\cos(0) = 1$ , o que resulta em uma interferência construtiva, resultando em um aumento da amplitude para  $2A$  (Figura 4). Se  $\phi = \pi$ , então  $\cos(\frac{\pi}{2}) = 0$ , o que resulta na interferência destrutiva (anulação da onda), isto é, a amplitude é 0 (Figura 5). A anulação da onda ocorrerá somente nos instantes que as fases das ondas interferentes permanecerem opostas.

Segundo Flo Menezes em [9], o som senoidal ou não existe na natureza, ou na natureza é só ele que existe de fato, uma vez que os demais sons (não-senoidais) são resultado de sobreposições de sons senoidais.

Para facilitar o estudo do som torna-se útil conceber a existência do som senoidal, quase toda fundamentação que se segue tem como base de sustentação a existência desse som desprovido de harmônicos, sem as ondas senoidais muitos resultados aqui mostrados não seriam obtidos facilmente, como por exemplo, a descrição da interferência de ondas.

A interferência é típica em casos de ondas que se propagam em diferentes direções e se interceptam em alguma região do espaço, e a superposição de ondas ocorre em um ponto em cada instante. Sendo as ondas sonoras tridimensionais (aproximadamente esféricas), isto é, ondas que se propagam em todas as direções, quando ocorre de duas fontes sonoras puntiformes produzirem simultaneamente sons, independente da distância que essas fontes estejam do receptor (por exemplo, um ouvinte), em qualquer ponto do espaço audível ocorrerá alguma interferência, ou seja, o ouvinte perceberá a interferência das ondas sonoras em qualquer posição que esteja em relação as fontes sonoras. Considerando o som como resultante da soma de ondas sonoras senoidais, a interferência dessas ondas em um ponto do espaço de propagação pode ser construtiva (Figura 4), quando ocorre o aumento da amplitude, ou destrutiva, quando ocorre um decréscimo da amplitude. Se as fases das ondas forem opostas e com mesma frequência, na superposição ocorre uma interferência destrutiva por completo (Figura 5), onde acontece das ondas se anularem e não produzirem um fenômeno sonoro [6, 9].

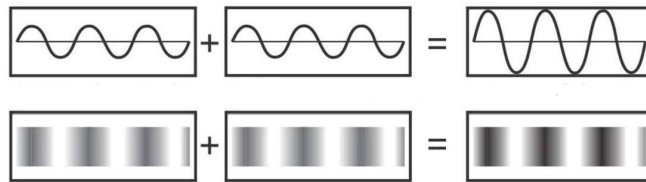


Figura 4: Interferência de ondas construtiva gerada pela superposição de duas ondas de mesma fase e frequência. Fonte: adaptado de [32].



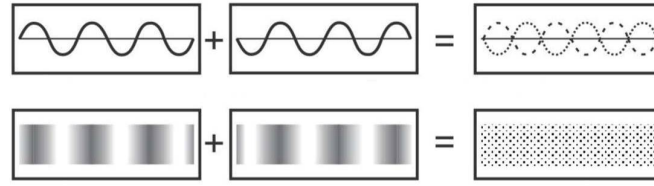


Figura 5: Interferência de ondas destrutiva gerada pela superposição de duas ondas de fases opostas e mesma frequência. Fonte: adaptado de [32].

Considerou-se para estes resultados ondas que se propagam no mesmo sentido, entretanto, a interferência de ondas sonoras acontece, geralmente, em diferentes sentidos de propagação, e os diferentes percursos implicam na diferença de fase de cada onda [6].

### 2.1.11 Timbre

O timbre não constitui um parâmetro do som, representa o resultado da combinação de vários atributos do som. Segundo a definição de Helmholtz, timbre é uma propriedade que permite a distinção de um som de um outro contendo a mesma altura e intensidade [9]. O timbre é um resultado da inter-relação entre diversos atributos como as alturas, as amplitudes, as durações e os comportamentos dinâmicos das amplitudes das parciais constituintes. O timbre é portanto um elemento composto, cada som que não seja senoidal ou ruído branco<sup>7</sup>, possui um timbre que o caracteriza por um conjunto de frequências do espectro sonoro [9].

## 2.2 Percepção do som (Psicoacústica)

A psicoacústica é uma área da psicofísica que aborda a relação entre estímulos sonoros e as sensações auditivas provocadas por esses estímulos, e é dividida em psicoacústica externa e interna. A psicoacústica externa trata da quantificação das sensações auditivas estabelecendo relações qualitativas entre os estímulos acústicos e as sensações auditivas. São objetos de estudo da psicoacústica, por exemplo, as sensações de volume sonoro, de tonalidade, de altura, dentre outras sensações obtidas através das características físicas do estímulo sonoro. Os métodos de medição utilizados no estudo da psicoacústica são análises psicológicas de quantificação das sensações auditivas a partir dos relatos do ouvinte ou de reações psicológicas verificadas por manifestações de ondas cerebrais. A psicoacústica interna trata dos mecanismos fisiológicos responsáveis pela transformação do estímulo sonoro em sensações auditivas. Resumidamente, pode-se dizer que

<sup>7</sup>Ruído branco é a combinação simultânea de sons senoidais com todas as frequências audíveis [9].



existe uma diferença entre o fenômeno sonoro e a percepção que temos do mesmo, a psicoacústica procura explicar essa diferença [36].

### 2.2.1 A audição humana

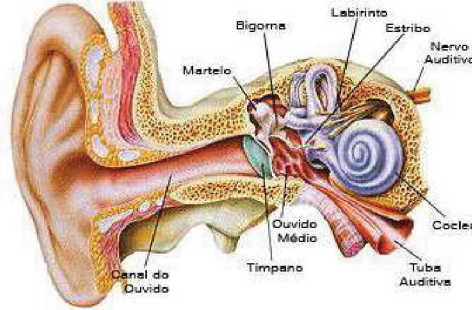


Figura 6: Orelha humana. Fonte: [8].

Quando as ondas sonoras chegam ao canal auditivo da orelha encontram o tímpano, que é uma membrana fina e sensível capaz de detectar pequenas variações de pressão. O tímpano transmite as vibrações para três ossos chamados martelo, bigorna e estribo. Durante a transmissão por esses ossos o som é ampliado atingindo a orelha interna que, devido ao seu formato de caracol em conjunto com os pêlos e líquido presente, faz os sinais sonoros se direcionarem até as células nervosas que, ao serem estimuladas, encaminham esses sinais pelo nervo auditivo até o cérebro humano (Figura 6). Após sofrer várias modificações, esses sinais fazem com que se tenha a percepção do som [4].

### 2.2.2 Relação qualitativa entre sensação e estímulo

Ernst Heinrich Weber, (1795 – 1878), estabeleceu que o aumento da sensação que produz um estímulo  $\Delta E$ , é diretamente proporcional ao incremento mínimo para se perceber o estímulo  $\Delta R$ , e inversamente proporcional ao estímulo inicial a ser comparado  $R$ , isto é,

$$\Delta E = k \frac{\Delta R}{R}, \quad (10)$$

sendo  $k$  a constante de proporcionalidade. A relação que existe entre a sensação ( $E$ ) com o estímulo ( $R$ ), isto é,  $E = E(R)$ , pode ser obtida pela Equação 10. Denotando  $dE$  e  $dR$  variações infinitesimais de  $R$  e  $E$ , respectivamente, pode-se reescrever a Equação 10 como

$$dE = k \frac{dR}{R}. \quad (11)$$

$$\Rightarrow \int dE = \int k \frac{dR}{R}$$

$$\Rightarrow E = k \ln(R/R_0)$$

onde  $E_0$  é a sensação umbral (Para  $R_0$  temos  $E_0=0$ ),  $E$  é a sensação (resposta ao estímulo),  $R_0$  o estímulo umbral (sensibilidade mínima),  $R$  é o Estímulo. Sendo  $\ln(x) \approx 2,3 \log(x)$ , tem-se

$$E \approx 2,3k \log(R/R_0). \quad (12)$$

A Figura 7 é um exemplo da relação entre o estímulo e a sensação.

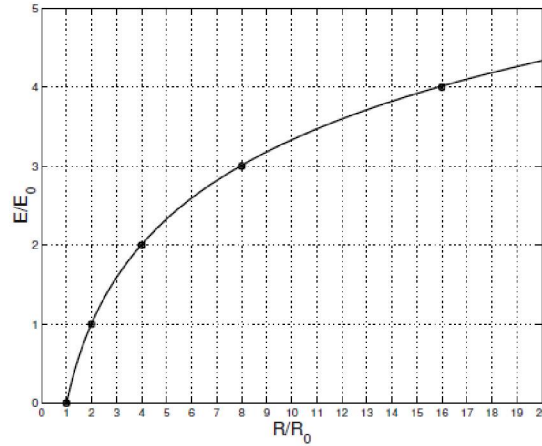


Figura 7: Relação qualitativa entre estímulo e sensação. Fonte: [5] p.5.

A Equação 12 representa a lei de Weber-Fechner, uma complementação da lei de Weber feita por Gustav Theodor Fechner, (1801 – 1887). Para ondas sonoras, conforme a lei de Weber-Fechner, temos que a intensidade da sensação sonora cresce proporcionalmente ao logaritmo da intensidade sonora (estímulo) [5], isto é, a sensação psicológica de sonoridade, que chamaremos de volume do som, varia aproximadamente com o logaritmo da intensidade e não com a própria intensidade, e a medida de volume sonoro obtida pelo logaritmo da intensidade sonora será denotada como nível de intensidade sonora.

### 2.2.3 Nível de intensidade sonora

O ouvido humano é sensível a pequenas variações de pressão, respondendo assim a ação de uma grande variedade de intensidades sonoras. Para facilitar a leitura das pressões sonoras ou das intensidades sonoras, convencionou-se a fazê-la através de níveis de pressão ou intensidade sonora<sup>8</sup>. Considerando que a sensação psicológica

<sup>8</sup>A unidade internacional do nível de pressão sonora é o decibel (dB), o *bel* ( $B$ ) foi inventado por engenheiros do Bell Labs (laboratório de pesquisa e de desenvolvimento da American Telephone and Telegraph dos Estados Unidos), em homenagem ao fundador do laboratório Alexander Graham Bell [33].

do som varia aproximadamente com o logaritmo do estímulo (Equação 12), ou seja, com o logaritmo da intensidade sonora, é, portanto, conveniente usar uma escala logarítmica para descrever esses níveis. O limiar de audibilidade depende tanto da amplitude como da frequência das ondas sonoras, pois a percepção do volume sonoro varia também com a frequência, como será visto na Seção 2.2.6 na Figura 10. Por exemplo, para uma frequência de  $1000\text{ Hz}$ , a intensidade mínima para causar uma sensação sonora é de  $I_0 = 10^{-12}\text{ W/m}^2$ , e se tomarmos  $I_0$  como referência, teremos como medida de nível de intensidade sonora uma escala logarítmica conhecida como *bel*, onde  $\log \frac{I}{I_0}$  é a medida em *bel* do nível de uma intensidade sonora  $I$ . Para facilitar a medição do nível de intensidade sonora convencionou-se a usar como padrão o decibel (*dB*), sendo  $1\text{ dB} = 0,1\text{ bel}$ , assim, o nível de intensidade sonora, *NIS*, fica definido pela equação

$$NIS = 10 \log \frac{I}{I_0}. \quad (13)$$

Se  $I = I_0$ , então  $NIS = 0\text{ dB}$ . O limite da dor, ou máxima intensidade sonora que o ouvido suporta é de  $I = 1\text{ W/m}^2$  para uma frequência de  $1000\text{ Hz}$ , que equivale a  $NIS = 120\text{ dB}$ . Isto significa que o ouvido suporta uma variação aproximada de  $120\text{ dB}$  de *NIS* [13].

Podemos determinar o nível de intensidade sonora em função da variação pressão sonora  $p$ , sendo a intensidade proporcional ao quadrado da variação da pressão, o que nos dá  $\frac{I}{I_0} = \frac{p^2}{p_0^2}$ , substituindo na Equação 13, obtemos a equação  $NIS = 10 \log \frac{p^2}{p_0^2} [dB]$ , denominada também como Nível de Pressão Sonora (*NPS*), temos assim a Equação 14,

$$NPS = 20 \log \frac{p}{p_0}. \quad (14)$$

A pressão sonora  $p_0$  usada como referência em geral é a mínima necessária para causar uma sensação sonora definida pelo “limiar de audibilidade”, onde  $p_0 = 20 \times 10^{-5}\text{ N/m}^2$  [14] [5]. Uma forma de quantificar a intensidade do som emitido por uma fonte sonora independente da sua localização, da distância entre a fonte sonora e o receptor, e independente do meio em que se propaga é através da potência sonora, isto é, da energia acústica emitida por uma fonte sonora por unidade de tempo, medida em watt ou em Joule por segundo, sendo  $1\text{ W} = 1\text{ J/s}$ . O Nível de Potência Sonora (*NWS*) medido em *dB* é em geral maior que os demais níveis, sendo a potência sonora usada como referência  $W_0 = 10^{-12}\text{ [W]}$ , temos assim o nível de potência sonora dado por  $NWS = 10 \log \frac{W}{W_0}$  [15]. Esse nível, o *NWS*, é geralmente usado em potenciômetros de amplificadores comerciais, porém, adotando como referência  $W_0$  a potência sonora máxima suportada pelo aparelho de som. Dessa forma, o *NWS* terá como valor máximo  $0\text{ dB}$ , e os demais valores negativos, em *dB*, representando as frações do nível de potência sonora máximo. Por exemplo,  $-3\text{ dB}$  indica, aproximadamente, metade

da potência máxima,  $-6 \text{ dB}$  representa aproximadamente um quarto do  $NWS$  máximo [16].

Os modelos matemáticos e computacionais apresentados neste trabalho adotam, para sinais de áudio, uma escala linear relativa a amplitude máxima  $x_0$  que é suportada pelo aparelho ou sistema de som. O sinal de áudio  $X$  pode representar a variação de tensão elétrica em volts, a variação de corrente elétrica ou a variação de pressão sonora, e mostra a variação de amplitude instantânea do sinal ao longo do tempo de áudio. Na escala linear a variação relativa do sinal ocorre no intervalo  $[-1, 1]$  e é dada por  $x/x_0$ , onde  $x$  representa a variação instantânea da grandeza física relacionada ao som. A opção pela escala linear é devido ao fato da representação gráfica em forma de onda dos dados numéricos do áudio ser mais adequada às características do sinal de áudio, principalmente quando o sinal for gerado por uma função matemática, como por exemplo, uma curva senoidal gerada por funções seno ou cosseno, nota-se que a senoide fica melhor caracterizada em uma escala linear em comparação, por exemplo, a uma escala logarítmica ( $\text{dB}$ ). Nas Figuras 8 e 9 são apresentados os gráficos de onda do sinal gerado pela função  $x(t) = 0,5\text{sen}(2\pi 400t)$ , com  $0 \leq t \leq 0,0025 \text{ s}$ , e os respectivos dados numéricos do áudio obtidos por amostragem e representados nas escalas linear ( $x/x_0$ ) e decibel ( $x_{\text{dB}} = 20 \log |x/x_0|$ ).

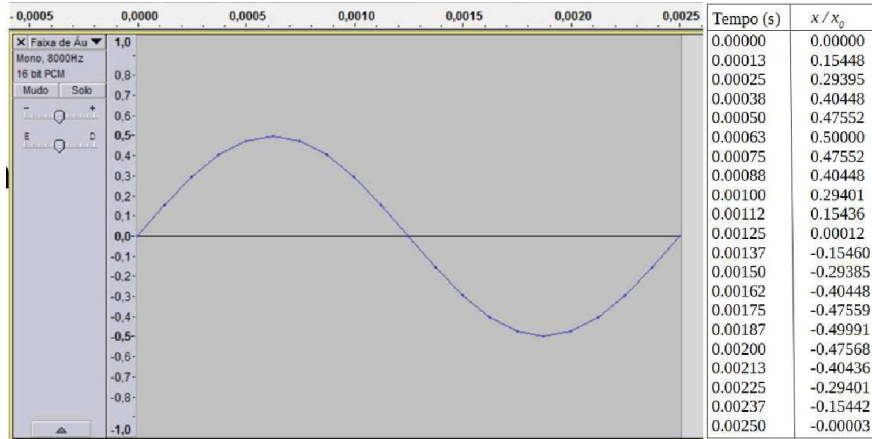


Figura 8: Escala linear. Fonte: adaptado da captura de tela do Audacity.



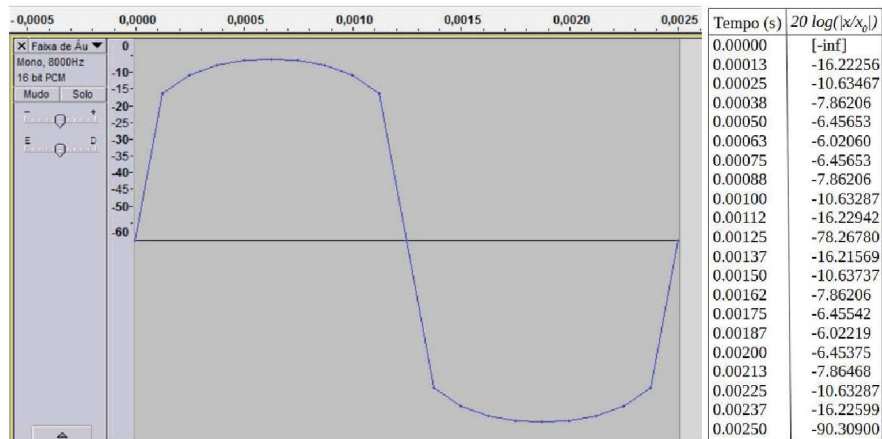


Figura 9: Escala decibel. Fonte: adaptado da captura de tela do Audacity.

### 2.2.4 Espectro sonoro

A transferência de energia através do processo oscilatório de compressão e rarefação determinada pela perturbação do meio material é o que ocasionam o surgimento das ondas sonoras. Essas perturbações podem ser periódicas, quando ocorrem em intervalos de tempo regulares, ou aperiódicas, quando são desprovidas de frequência. Quando existe um período regular de perturbação do meio em algum intervalo de tempo dizemos que existe uma frequência, isto é, existem repetições do ciclo oscilatório. O espectro sonoro é o conjunto de todas as frequências que se associam a todas as perturbações longitudinais de um meio material, podem ou não ser audíveis. O ouvido humano capta o som gerado com frequências de ondas mecânicas compreendidas na faixa entre  $20\text{ Hz}$  e  $20000\text{ Hz}$  aproximadamente. As ondas sonoras com frequência menor que  $20\text{ Hz}$  são conhecidas como infrassônicas, e acima de  $20000\text{ Hz}$  são chamadas de ultrassônicas.

### 2.2.5 Altura

A altura é uma propriedade perceptiva dos sons gerados por ondas periódicas, a altura é um parâmetro que associa a frequência da onda à percepção de som grave, médio ou agudo. Sons graves possuem frequências baixas e sons agudos possuem frequências altas, esta relação entre altura e frequência foi mostrada pela primeira vez por Robert Hooke (1635-1703), ao colocar um cartão na frente a uma roda dentada que, ao girar, os dentes batiam no cartão emitindo um determinado som associado a uma nota musical. Hooke observou que ao aumentar a velocidade de rotação da roda a frequência aumentava proporcionalmente. Sabendo da velocidade de rotação e o número de dentes da roda, foi possível calcular a frequência dos sons produzidos [2, 34]. Mesmo sendo inegável a relação direta entre altura e frequência, a percepção de grave, médio ou agudo se dá de maneira subjetiva em cada sujeito, podendo até mesmo em um único sujeito provocar uma



percepção diferente de altura de um para o outro ouvido. Em geral, cada pessoa possui uma própria percepção de altura [2, 9].

### 2.2.6 Curvas isofônicas

Para a percepção do volume sonoro deve-se considerar a sensibilidade do ouvido humano, que tem respostas diferentes para ondas sonoras com diferentes frequências. Para que uma onda sonora seja percebida, além de estar na faixa de frequências audíveis, requer também subjetivamente uma intensidade sonora mínima. Esta intensidade não é a mesma para todas as frequências, o que pode ser verificado com certos testes auditivos. Por exemplo, ao diminuir o volume de um aparelho de som temos a impressão de que o som fica mais agudo, pois o ouvido humano, em geral, precisa de um nível de pressão sonora menor para captar ondas com altas frequências em relação as ondas com frequências mais baixas. Sendo em geral o som composto de frequências altas, médias e baixas, quando o volume do aparelho de som é diminuído, o ouvido deixa de perceber essas frequências mais baixas, tendo a impressão de que o som ficou menos grave, ou seja, mais agudo. Essa percepção pode ser representada qualitativamente num gráfico de “Curvas Isofônicas”, através da referência *phon*<sup>9</sup> obtido de forma experimental. Na Figura 10, o eixo horizontal do plano representa a escala de frequências em *Hz*, o eixo vertical representa a escala dos níveis de pressão sonora em *dB*, e as linhas presentes no plano representam as relações entre frequências e níveis de pressão sonora dada pela referência *phon*, sendo cada linha indicada com a intensidade em *dB* do tom puro usado como referência. Por exemplo, a primeira linha, de baixo pra cima, é relativa a um tom puro de referência com 0 *dB* (0 *phon*), e a última linha, na parte de cima do plano, é a relação entre nível de pressão sonora e frequência dada por um tom puro de referência com 90 *dB* (90 *phon*). A pessoa que se submete ao teste auditivo escuta um tom puro de 1000 *Hz* como referência, e depois muda para outra frequência, ajustando o nível de intensidade até que se perceba um volume igual ao tom de referência [5].

---

<sup>9</sup>O *phon* é uma unidade de medida de intensidade que corresponde ao nível de som, em *dB*, de um tom puro de 1000 *Hz*, que produz o nível de volume igual ao nível de intensidade desejado [5].

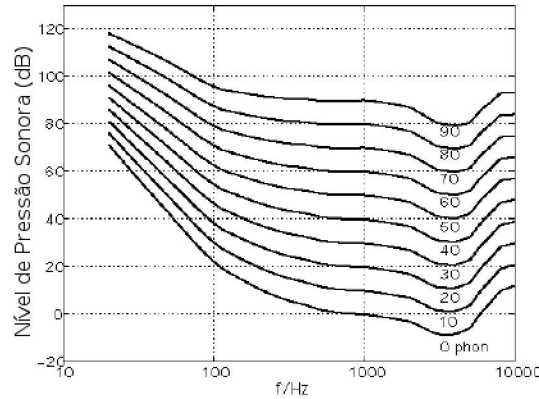


Figura 10: Curvas isofônicas. Fonte: adaptado de [5].

Por exemplo, para um tom puro de referência de  $1000\text{ Hz}$ , com nível de pressão sonora de  $20\text{ dB}$  ( $20\text{ phon}$ ), é necessário um volume de  $40\text{ dB}$ , aproximadamente, em um som com frequência de  $100\text{ Hz}$ , para percebê-lo com o mesmo volume.

## 2.3 Análise espectral

O som mais presente em nosso cotidiano é o definido como complexo. Um som complexo é um som constituído por mais de uma frequência, formado por parciais harmônicas e não harmônicas, e o conjunto de todas as parciais é conhecido como espectro sonoro.

O espectro sonoro pode ser representado matematicamente como uma série trigonométrica, conhecida como série de Fourier<sup>10</sup>, parte integrante da análise de Fourier, relacionada a fenômenos periódicos e não periódicos. Uma onda sonora complexa e periódica qualquer é chamada de onda composta, e pode ser decomposta em uma série de ondas senoidais com múltiplas frequências inteiras a partir de uma certa frequência fundamental. As frequências parciais dessa série são chamadas harmônicas, e a frequência de cada uma dessas harmônicas possui uma amplitude máxima  $A$  determinada ( $A = \Delta P_m$ ). A forma espectral de uma onda sonora é um gráfico das frequências e amplitudes máximas,  $f \times A$ , das componentes parciais (senoides) da função  $v$  que representa a onda. Essa forma é uma alternativa à representação na forma de onda, que é um gráfico que representa as amplitudes instantâneas  $v(t)$  da onda sonora no instante  $t$ . A Figura 11 ilustra essas duas possibilidades nos planos  $t \times v$  e  $f \times A$ , respectivamente. O modelo temporal (forma de onda) descreve uma grandeza física (variação de pressão) ao longo do tempo, e pode ser representado por funções reais com domínio real (tempo); e o modelo espectral (forma espectral) analisa as densidades de

<sup>10</sup>Jean Baptiste Joseph Fourier, Físico-Matemático francês (1768 - 1830), utilizou séries trigonométricas em sua obra, "*Théorie Analytique de la Chaleur*", para estudar os fenômenos de condução de calor [35].

frequências, no domínio das frequências, através das séries de Fourier e da análise espectral possibilitada pela transformada discreta de Fourier [21, 26].

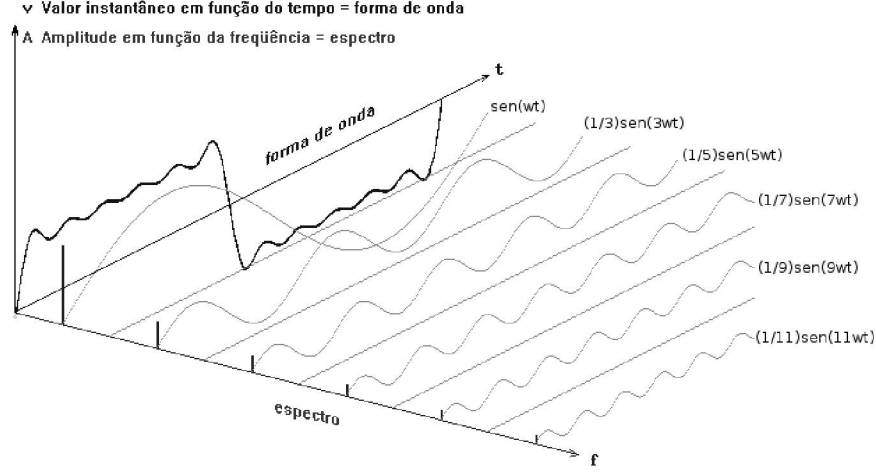


Figura 11: Relação entre forma de onda e espectro. Fonte: adaptado de [19].

Na Figura 11, tem-se  $v(t) = \sum_{k=1}^6 \frac{\text{sen}((2k-1)\omega t)}{2k-1}$ , sendo  $\frac{\text{sen}((2k-1)\omega t)}{2k-1}$  a frequência fundamental se  $k = 1$ , e representando cada harmônico se  $k \neq 1$ . Cada parcela dessa soma é denominada componente parcial de  $v(t)$  ou frequência parcial do espectro da função  $v$ . Se  $k$  for para o infinito,  $v(t)$  representará uma onda quadrada.

Uma função  $v(t)$  é dita periódica se existir um período  $\tau$ , tal que  $v(t) = v(t+\tau)$  para todo  $t$  real. As funções periódicas mais conhecidas são as funções seno e cosseno, com elas pode-se construir, para uma dada função  $v(t)$ , a série

$$a_0 + \sum_{n=1}^{\infty} a_n \cos(n\omega_0 t) + b_n \text{sen}(n\omega_0 t), \quad (15)$$

que é conhecida como a série de Fourier de  $v(t)$  [19, 24, 35]. Esta série é periódica com período fundamental  $\tau$  e frequência fundamental  $\omega_0 = \frac{2\pi}{\tau}$ , e os coeficientes  $a_0$ ,  $a_n$  e  $b_n$ , chamados de coeficientes de Fourier, são determinados pelas equações

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} v(t) dt, \quad (16)$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} v(t) \cos(n\omega_0 t) dt, \quad (17)$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} v(t) \text{sen}(n\omega_0 t) dt, \quad (18)$$

$n = 1, 2, 3, \dots$

A série de Fourier de  $v(t)$  pode também ser escrita na forma complexa, escrevendo  $\text{sen}(n\omega_0 t) = \frac{e^{iz} - e^{-iz}}{2i}$  e  $\text{cos}(n\omega_0 t) = \frac{e^{iz} + e^{-iz}}{2}$  e simplificando<sup>11</sup>, obtém-se, assim, a expansão em Série de Fourier Complexa (ou Exponencial) de  $v$  dada por,

$$v(t) \approx \sum_{n \in \mathbb{Z}} c_n e^{in\omega_0 t}.$$

A partir da representação da série de Fourier na forma complexa obtém-se a transformada de Fourier de  $v$ , definida por  $V$ , que associa a cada função absolutamente integrável  $v : \mathbb{R} \rightarrow \mathbb{R}$  a função  $\hat{v} : \mathbb{R} \rightarrow \mathbb{C}$  definida pela equação  $\hat{v}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} v(t) e^{-i\omega t} dt$ , onde, a inversa  $V^{-1}$ , chamada transformada de Fourier inversa, associa cada função  $\hat{v} : \mathbb{R} \rightarrow \mathbb{C}$  pertencente ao conjunto imagem de  $V$ , à função integrável  $v : \mathbb{R} \rightarrow \mathbb{R}$  definida por  $v(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{v}(\omega) e^{i\omega t} d\omega$ , quando  $v$  é contínua, tem-se  $V^{-1}(V(v)) = v$ . Essa é a grande contribuição de Fourier para matemática moderna, a análise do espectro de frequências relacionando um modelo temporal com um modelo espectral<sup>12</sup>, é pela transformada de Fourier que se pode descobrir as senoides que precisam ser adicionadas para criar a forma da onda resultante, obtendo exatamente as frequências das ondas a ser somadas, e em quais quantidades, para a representação do sinal exato de uma onda resultante [26, 19, 9, 27, 35].

## 2.4 Processamento de sinais de áudio

O processamento de sinais de áudio é o ato de extrair e manipular informações contidas em um áudio. Depois de transduzido para um sinal elétrico, o processamento do sinal de áudio pode ser feito de forma analógica, com sinal de tempo contínuo, ou de forma digital, com sinal de tempo discreto. Os sinais de tempo discreto são bem mais versáteis, pois o processamento digital tem um custo computacional bem menor. Os sinais de áudio transportam ou armazenam informações de variação de pressão e podem ser representados em função de uma ou mais variáveis.

A forma mais comum de representar esses sinais é através da variação da pressão em função do tempo. A gravação de sinais de áudio teve origem com Thomas Edison, em 1877, utilizando-se de um processo mecânico onde o som, ao vibrar um diafragma, transmitia a vibração da onda para uma agulha fixa no mesmo, a agulha riscava um cilindro em movimento de rotação e translação produzindo um sulco que imprimia de forma contínua uma inscrição do som. A revolução na captura de sinais de áudio iniciou-se com Lee De Forest, em 1915, ao inventar um amplificador de tubo a vácuo, a partir daí a gravação deixava de ser apenas acústica e passava a ser também elétrica [37, 38].

<sup>11</sup>Demonstrações e demais detalhes desta seção podem ser encontrados em [26]

<sup>12</sup>Demonstrações e demais detalhes em [27].



Desde então o processo evoluiu, surgiram os microfones, que são hoje os transdutores mais comuns usados para esse fim. Existem vários tipos de microfones, que se diferenciam pela forma como é feita a captura de sinais acústicos e a conversão em sinais elétricos, sendo mais empregados atualmente dois tipos de microfones, os dinâmicos de bobina móvel e os capacitivos.

Os microfones dinâmicos de bobina móvel funcionam através de uma pequena bobina, fixa em um diafragma, que é posicionada em um campo magnético de um ímã permanente. O som ao ser direcionado para o microfone, com uma certa intensidade, faz com que as ondas sonoras vibrem o diafragma, fazendo a bobina se mover no campo magnético, gerando uma tensão variável nos terminais da bobina, resultando em uma tensão induzida que provoca uma corrente elétrica diretamente relacionada a variação do campo magnético. Os diafragmas são formados por membranas específicas para uma determinada faixa de frequências, assim, para cada tipo de membrana, tem-se uma resposta melhor dependendo da altura do som (grave, médio ou agudo). Por isso existem tipos distintos de microfones dinâmicos para diferentes aplicações sonoras, existindo também microfones dinâmicos versáteis, que combinam múltiplas membranas para abranger diferentes faixas de frequências do espectro do áudio para a captação mais fidedigna do sinal.

Um outro tipo de microfone bastante empregado é o capacitivo, conhecido como microfone condensador, que utiliza de princípios da eletrostática. Independente do tipo de microfone, todos se baseiam no mesmo princípio, transformar um sinal de ondas de pressão em um sinal elétrico [42, 37].

Podemos classificar um sinal de áudio entre determinístico ou aleatório, e entre contínuo ou discreto. Um sinal é classificado como determinístico quando os valores instantâneos podem ser previstos no tempo, são sinais que podem ser expressos por funções matemáticas com domínio no tempo, podem ser preditos antes de ocorrerem, um sinal gerado por uma função senoidal como, por exemplo,  $x(t) = 0,2\sin(2\pi 100t)$  é um exemplo de sinal determinístico. Um sinal é classificado como aleatório quando existe incerteza sobre seus valores antes da ocorrência, são geralmente representados por suas características estocásticas, como média, variância, desvio padrão, autocorrelação, etc. Um sinal de áudio gerado por uma gravação aleatória de voz, ou por um ruído, são exemplos de sinais aleatórios [39]. Neste trabalho abordaremos tanto os sinais de áudio aleatórios como os determinísticos.

Um sinal de áudio de tempo contínuo é um sinal que está definido em qualquer instante do tempo. Na obtenção desse sinal, as variações de pressão de uma onda sonora são transduzidas (transformadas) para um sinal elétrico contínuo denominado como analógico, ou seja, um sinal  $x(t)$  de tempo contínuo definido para todo instante  $t$ . De modo geral, um sinal se diz contínuo se o seu domínio for  $\mathbb{R}$  ou um intervalo de  $\mathbb{R}$ . O sinal contínuo, no entanto, admite descontinuidades, isto é, um sinal contínuo nem sempre é uma função contínua. Para representar um sinal contínuo no tempo são necessários infinitos pontos, o que dificulta o pro-

cessamento computacional do sinal, devido a limitação de memória dos recursos computacionais. Para facilitar esse processamento é que surgiu a conversão de um sinal contínuo para um sinal discreto [39, 21].

Um sinal de áudio de tempo discreto é um sinal que é obtido a partir de amostras do sinal de tempo contínuo. Um sinal discreto tem como domínio o conjunto dos  $\mathbb{Z}$ , ou um intervalo de  $\mathbb{Z}$ , isto é, dado um sinal contínuo  $x(t)$  de duração finita,  $t \in [t_0, t_f]$ , e definido um intervalo de tempo  $T$ ,  $T \leq t_f$ , pode-se converter esse sinal contínuo para um sinal discreto  $x(nT)$ , com  $n \in \mathbb{Z}$ , onde  $nT$  define os valores ou amostras igualmente espaçadas do sinal contínuo, sendo o sinal de áudio discreto representado pela sequência  $x[n] = x(nT)$ . Os sinais de tempo discreto com duração finita podem, portanto, ser representados por finitos pontos, tanto no domínio do tempo como no das frequências, o que torna viável o seu processamento digital. Na Figura 12 temos uma representação gráfica da passagem de um sinal contínuo para um sinal discreto, o amostrador representa a chave do circuito que fragmenta o sinal contínuo em pequenos intervalos de tempo igualmente espaçados, coletando, durante o breve instante que a chave permanece aberta, os valores aproximados das amostras [21].

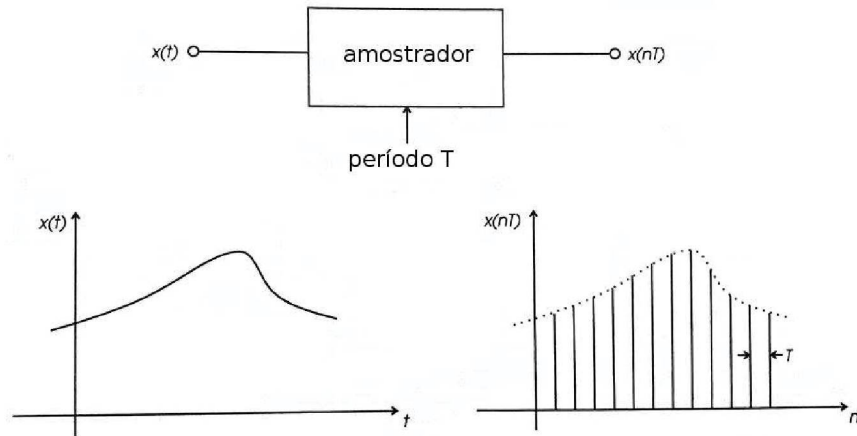


Figura 12: Amostragem de um sinal de tempo contínuo  $x(t)$  para obtenção de um conjunto discreto  $x(nT)$ . Fonte: adaptado de [21].

Amostragem é o procedimento pelo qual um sinal contínuo é transformado em um sinal discreto. Amostrar um sinal contínuo é recolher valores desse sinal em instantes discretos. Geralmente o intervalo de tempo entre as amostras é constante, e nesse caso é definido como período de amostragem ( $T$ ). Quando as amostras são coletadas em intervalos regulares, a quantidade de vezes em que se coletam amostras de um sinal contínuo em uma unidade de tempo é conhecida como taxa de amostragem ( $F_s$ ), normalmente medida em Hertz ( $Hz$ ), isto é, um número inteiro de amostras por segundo. Quanto maior for a taxa de amostragem, mais fidedigno será o sinal em relação ao sinal contínuo original.

O que permite amostrar um sinal é um circuito com uma chave que se fecha em intervalos regulares de tempo, com o período desse intervalo definido pela taxa de amostragem ( $Fs$ ). Por exemplo, para uma taxa de amostragem de 1000  $Hz$  ( $Fs = 1000$  amostras por segundo), a chave se fecha a cada 1  $ms$ , isto é,  $T = 1\ ms$ . Ao se fechar por um breve período, o sinal resultante ocorrerá em forma de pulsos curtos, chamados Pulsos Modulados em Amplitudes ( $PAM$ ), com amplitude bem próxima do valor instantâneo do sinal [22].

Para exemplificar esse processo, considere um sinal contínuo representado pela função  $x(t) = 0,5\text{sen}(2\pi 400t)$ , com  $0 \leq t \leq 0,0025$  segundos. Para discretizar o tempo  $t$ , vamos aplicar uma taxa de amostragem  $Fs = 8000\ Hz$ , definindo assim um período de amostragem  $T = \frac{1}{Fs} = 0,000125$  segundos, isto é, vamos obter uma amostra a cada 0,125  $ms$ , como se pode ver na Tabela 2. A relação entre  $T$  e  $Fs$ , dada por

$$T = \frac{1}{Fs}, \quad (19)$$

é fundamental para os modelos desenvolvidos neste trabalho.

$n$	$nT$	$x[n] = x(nT)$
1	0.000125	0.1545084971875
2	0.00025	0.2938926261462
3	0.000375	0.4045084971875
4	0.0005	0.4755282581476
5	0.000625	0.5
6	0.00075	0.4755282581476
7	0.000875	0.4045084971875
8	0.001	0.2938926261462
9	0.001125	0.1545084971875
10	0.00125	0
11	0.001375	-0.1545084971875
12	0.0015	-0.2938926261462
13	0.001625	-0.4045084971875
14	0.00175	-0.4755282581476
15	0.001875	-0.5
16	0.002	-0.4755282581476
17	0.002125	-0.4045084971875
18	0.00225	-0.2938926261462
19	0.002375	-0.1545084971875
20	0.0025	0

Tabela 2: Amostragem para o sinal  $x(t) = 0,5\text{sen}(2\pi 400t)$ , com  $0 \leq t \leq 0,0025\ s$ .

A partir da Tabela 2 pode-se construir o gráfico  $nT \times x[n]$ , onde  $nT$  (valores no

eixo horizontal) representa o instante de coleta da  $n$ -ésima amostra, em segundos, e  $x[n]$  (valores no eixo vertical) a amplitude no instante  $nT$ , como apresentado na Figura 13.

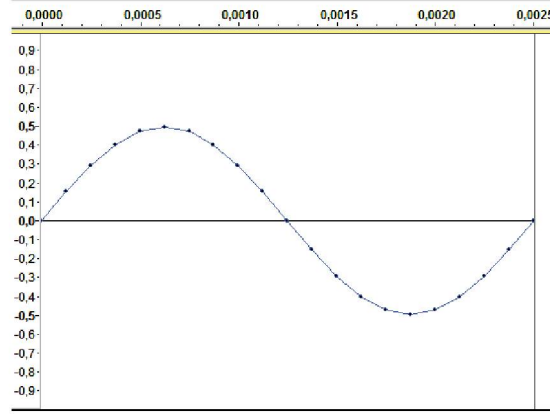


Figura 13: Gráfico  $nT \times x[n]$  (pontos) da função  $x(t)$  (linha contínua) discretizada por um período de amostragem  $T = 0,125 \text{ ms}$ , amostras dadas pela Tabela 2. Fonte: adaptado da captura de tela do Audacity.

A discretização do tempo  $t$  é a primeira etapa da conversão de um sinal analógico para um sinal discreto. Para que no processo de reconstrução do sinal contínuo a partir de sinal discreto não haja perda de informação, deve-se ter o cuidado de obter uma quantidade mínima de amostras do sinal contínuo, quantidade essa que é definida pelo Teorema 1 (Teorema de Shannon-Whittaker), que é uma complementação do teorema desenvolvido pelo engenheiro sueco Harry Nyquist (1889 – 1976), ao analisar a reconstrução de um sinal discreto.

**Teorema 1 (de Shannon-Whittaker)** *Sejam  $x$  um sinal de banda limitada<sup>13</sup>, e  $\Omega$  o menor valor de frequência tal que  $\text{sup}(\hat{x}) \subset [-\Omega, \Omega]$ . O sinal  $x$  pode ser reconstruído de forma exata a partir de um conjunto uniforme de amostras  $\{nT; n \in \mathbb{Z}\}$  se  $T < \frac{1}{2\Omega}$  [28, 39].*

O valor  $\frac{1}{2\Omega}$  é chamado de “limite de Nyquist”, ou frequência de Nyquist. Pode-se notar que

$$T < \frac{1}{2\Omega} \Leftrightarrow \frac{1}{T} > 2\Omega \Leftrightarrow Fs > 2\omega.$$

Logo, ao tentar reconstruir um sinal com um período de amostragem maior que a frequência de Nyquist, ou seja, se a taxa de amostragem for menor que o dobro da

<sup>13</sup>“Banda” é um termo técnico muito usado em engenharia de telecomunicações para se referir a diferença entre duas frequências. “Largura de banda” é a diferença entre a maior e menor frequência parcial do sinal, e “banda limitada” é um termo usado para indicar que a largura de banda é finita.



maior frequência parcial, ocorre um efeito chamado aliasing<sup>14</sup>, ocorrendo perda de informação, como ilustrado na Figura 14. Também é necessário que o sinal seja de banda limitada, isto é, que o sinal contínuo tenha espectro com frequências angulares definidas em um intervalo limitado, para ser possível sua reconstrução exata a partir do sinal discreto resultante, caso contrário o seu espectro será certamente replicado com sobreposição, ocasionando perda de informação.

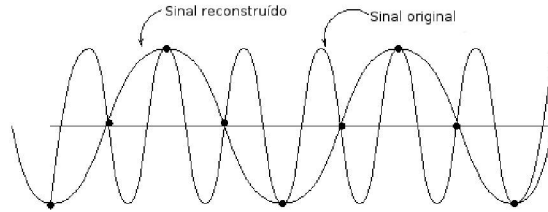


Figura 14: Aliasing, erro cometido ao usar uma taxa de amostragem menor que a definida pelo Teorema 1. Fonte: adaptado de [25].

A taxa de amostragem padrão para sinais de áudio digital é de  $44100\text{ Hz}$ , pois, sendo a frequência máxima audível ao ouvido humano de  $20\text{ KHz}$ , então, basta amostrar a  $40\text{ KHz}$  (o dobro da frequência máxima) para contemplar todo o espectro sonoro audível, conforme o Teorema 1. O valor  $44100\text{ Hz}$  é usado como taxa de amostragem padrão, pois ocorre um acréscimo de tolerância e por ter como fator o produto dos quatro primeiros números primos  $((2 \times 3 \times 5 \times 7)^2 = 44100)$ , para facilitar o fracionamento do conjunto de dados amostrados. Com essa taxa de amostragem temos uma forma de remover a maioria das frequências parciais acima de  $20\text{ KHz}$ .

A segunda etapa da conversão do sinal analógico para um sinal discreto é a quantização, que é a discretização das amplitudes instantâneas do sinal. Essa quantização em valores finitos de *bits* é feita através de um circuito conhecido como conversor analógico-digital A/D ou ADC. Nessa conversão, cada amostra é transformada em uma quantidade finita de *bits*, por exemplo, com  $8\text{ bits}$  é possível representar 256 valores distintos. Os valores reais das amplitudes amostradas são quantizados para um valor maior ou menor conforme o nível de decisão do projeto ADC, essa diferença entre o valor real e o quantizado é conhecida como erro de quantização, que provoca o surgimento de um sinal aleatório conhecido como ruído de quantização [20, 21].

Para transformar o sinal discretizado em um sinal digital é necessário, ainda, a codificação do sinal quantizado. Ela é feita por modulação por código de pulso (PCM). Depois de quantizado, o sinal é codificado para números binários para ser analisado por um processador digital de sinais, quanto maior a quantidade de

<sup>14</sup>Aliasing em português significa “frequências réplicas”, esse termo é usado para descrever o efeito de superposição dos espectros, ocasionando erro de reconstrução do sinal, onde frequências mais altas passam a ser menores [20].

*bits* utilizada na codificação melhor será a resolução numérica de cada amostra e menor será o erro de quantização. A Figura 15 apresenta a relação de um sinal digitalizado com um sinal analógico (contínuo), no eixo horizontal temos as  $n$  amostras discretizadas e no eixo vertical as amplitudes quantizadas e convertidas em sinal elétrico (*Volts*).

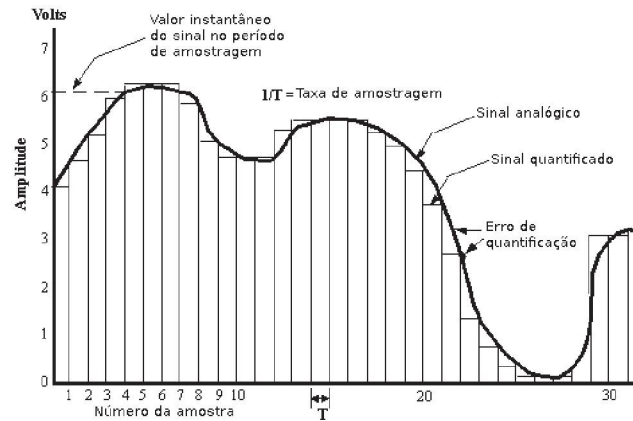


Figura 15: Quantização. Fonte: adaptado de [23].

A codificação é o passo seguinte, após o sinal ser quantizado é preciso fazer a codificação em seqüências de *bits*, isto é, transformar em um sinal digital binário. Como mostra a Figura 16, o sinal quantizado é modulado por um método conhecido como modulação de pulsos por código, PCM (Pulse Code Modulation), é a forma padrão para áudio digital. Cada pulso PAM de amplitude variável é transformado em uma seqüência de *bits* com amplitude fixa e com valores 0 ou 1, com um código que representa o valor do pulso PAM original, arredondado pelo erro de quantização.

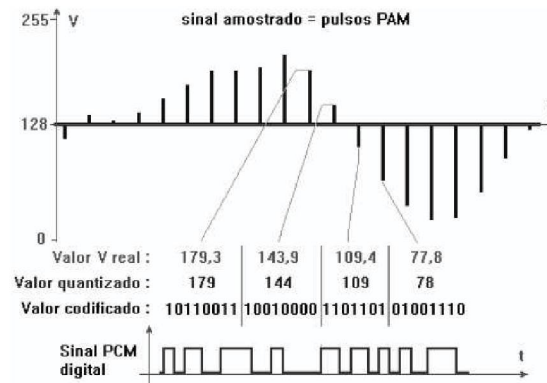


Figura 16: Sinal amostrado para modulação por código de pulso com 8 *bits* no formato .wav. Fonte: [22].

A quantização e a codificação são recursos já disponíveis nas plataformas computacionais usadas neste trabalho, a manipulação desses processos podem interferir, por exemplo, na resolução das amostras numéricas, e são, portanto, fundamentais para a obtenção dos dados a serem modelados matematicamente [22]. A discretização, a quantização e codificação, embora fundamentais para a realização das modelagens desenvolvidas neste trabalho, não serão explicitamente abordadas na construção dos mesmos. Assume-se aqui que o ponto de partida para os modelos é um áudio já digitalizado. Para a modelagem das ondas sonoras, o principal parâmetro que conduzirá a construção dos modelos será quantidade de amostras  $n$ , que, junto com o tempo de áudio  $t_f - t_0$  e a taxa de amostragem  $F_s$ , possibilita as implementações por manipulação matricial dos efeitos desejados.

Um processo importante para a avaliação das aplicações dos modelos deste trabalho consiste na transformação de um sinal discreto em um sinal contínuo, ou seja, fazer a conversão de um sinal digital em um sinal analógico. Na prática, usa-se para isso conversores digitais-analógicos (DAC), que são circuitos eletrônicos que transformam um sinal elétrico discreto em um sinal contínuo. As conversões dos sinais discretos para os sinais contínuos usadas em alguns modelos deste trabalho são feitas usando interpolação por funções splines lineares e cúbicas. Os gráficos apresentados na seção seguinte, obtidos no Scilab, Audacity e GeoGebra, representam a conversão do sinal discreto (pontos) em um sinal contínuo (linhas do gráfico).

### 3 Modelagem matemática e computacional de efeitos em ondas sonoras

Nesta seção apresentam-se os modelos matemáticos e computacionais de alguns efeitos sonoros: o controle de volume, que é um parâmetro fundamental na definição do som; o envelopamento dinâmico, que representa a variação, natural ou não, da intensidade sonora no intervalo de tempo do som; delays, reverberações e ecos, que são efeitos que simulam alterações que podem ocorrer naturalmente em certos ambientes; variação da frequência de um som harmônico associado a uma nota musical, que é feita através da alteração da quantidade de amostras de um áudio digital; reconstrução sintetizada de um sinal e composição de sons, utilizando a representação gráfica (forma de onda) em um modelo temporal gerado por um som gravado ou de dados numéricos determinados por alguma função matemática.

Estabelece-se que o dado inicial para o processo de modelagem é um áudio digitalizado de curta duração. Este áudio consiste basicamente de um conjunto de dados numéricos dado por um modelo temporal e obtido por algum critério de amostragem. No modelo temporal esses dados são as amplitudes, obtidas geralmente a partir de uma taxa de amostragem (em Hertz), com a resolução de



cada amostra (em bits), e o intervalo de tempo do áudio analisado (em segundos).

Para modelar os efeitos propostos é necessário, primeiramente, ter acesso aos dados numéricos do arquivo de áudio. Denomina-se esta etapa inicial por captura do sinal de áudio. Esta captura pode ser feita de diversas maneiras. Por exemplo, a partir de um arquivo de som qualquer já existente e com uma das extensões .wav, .mp3 ou .au, pode-se escolher uma das seguintes alternativas para obter os dados numéricos a ele associados: (I) usar o software Audacity para obter os dados em um arquivo com extensão .txt, .csv ou .html, o qual deverá posteriormente ser aberto e lido por um outro software adequado (bloco de notas, planilha eletrônica, GeoGebra, Scilab, etc), que neste trabalho será o Scilab; (II) Usar diretamente o software Scilab, que importará os dados numéricos já na forma adequada aos modelos aqui apresentados.

No caso da escolha da alternativa (I), apresenta-se no Anexo A um tutorial que descreve passo a passo como proceder no Audacity. Além disso, no Anexo B, Modo 1, apresenta-se um tutorial que mostra como importar esses dados no Scilab, apenas para os arquivos com extensões .txt ou .csv. Deste modo, recomenda-se fortemente o uso de uma destas duas extensões para obter os dados numéricos.

No caso da escolha da alternativa (II), encontra-se no Anexo B, Modo 2, um tutorial apresentando vários códigos, cada um deles adequado para cada tipo de extensão do arquivo de áudio (.wav, .mp3 e .au) que farão a captura de todos os dados numéricos necessários.

Além disso, há ainda a opção de se gravar um áudio de interesse no próprio Audacity, onde se pode definir o canal de gravação (esquerdo, direito ou ambos), a taxa de amostragem e resolução de cada amostra, como também gravar um tom puro ou outro tom sintetizado pelo próprio Audacity.

### 3.1 Representação matemática dos dados iniciais

Suponha agora que a captura de um sinal de áudio monoaural tenha sido realizada segundo alguma das alternativas descritas, (I) ou (II). Têm-se, portanto, duas possibilidades de dados importados: (a)  $n$  amostras  $x_i$ , com  $i \in I = \{i \in \mathbb{N}; 1 \leq i \leq n\}$ , de intensidades sonoras equiespaçadas no tempo, com a opção de obter  $n$  instantes de tempo  $t_i$  em que acontece as intensidades sonoras  $x_i$  correspondentes; (b)  $n$  amostras  $x_i$ , com  $i \in I = \{i \in \mathbb{N}; 1 \leq i \leq n\}$ , de intensidades sonoras equiespaçadas no tempo, a taxa de amostragem  $F_s$  com a qual essas amostras foram coletadas e o valor  $b$  da quantidade de bits de resolução de cada amostra. Em ambos os casos as intensidades sonoras são fornecidas. Elas deverão ser armazenadas em uma matriz linha  $X_{1 \times n} = [x_1 \ x_2 \ x_3 \ \dots \ x_i \ \dots \ x_n]$ . Nos dois casos é possível determinar os instantes  $t_i$  a partir da seguinte análise.

Sabe-se que a taxa de amostragem  $F_s$  é inversamente proporcional ao período



de amostragem  $T$ , isto é,

$$Fs = \frac{1}{T}$$

e  $T$  é aproximadamente o tempo de duração do som digital dividido pelo número de amostras, ou seja,  $T = \frac{t_n}{n} \approx \frac{t_f}{n}$ , onde  $t_f$  é o tempo de duração do som original. Portanto,  $t_n = nT = \frac{n}{Fs}$ , assim, cada elemento  $x_i$  representa uma amostra da amplitude sonora no instante  $t_i = \frac{i}{Fs}$ , e da mesma forma, temos  $i = Fs \cdot t_i$ . Em geral, neste trabalho os modelos de efeitos sonoros serão representados por uma função  $F$  que associa cada matriz  $X$  de amostras de intensidades sonoras uma matriz  $Y$  com intensidades sonoras modificadas. A modelagem a seguir descreverá matematicamente e computacionalmente a variação do volume de áudio.

## 3.2 Variação do volume do sinal de áudio

Mantendo as características do som como timbres, frequências e duração do áudio, a variação do volume sonoro consiste de uma transformação linear da forma  $Y = k \cdot X$ , onde  $X$  representa a matriz com as amostras das intensidades sonoras e  $k$  é um coeficiente numérico a ser definido para determinar a taxa de variação proporcional da amplitude de cada amostra. A matriz  $Y$ , de mesma ordem de  $X$ , representa as amostras do áudio transformado. Portanto, pode-se determinar a variação linear da amplitude de cada amostra simplesmente variando o valor de  $k$ , isto é, a variável  $k$  determina o volume do áudio.

Apresenta-se a seguir o algoritmo em código Scilab seguido da descrição de cada uma de suas linhas.

### 3.2.1 Algoritmo no Scilab para a variação do volume de áudio no formato .wav

```
1. clear; stacksize('max'); format(16);
2. N=input("Escreva o nome do arquivo de audio wav:", "string");
3. [X,Fs,bits]=wavread(N);
4. [nl,nc]=size(X);
5. c=strcat(['Escreva a faixa a ser modificada, valor entre 1 e ',
string(nl), '(para todas as faixas digite :)', ': ']);
6. d=input(c);
7. k=[X(d,:)];
8. i=int(100/(max(abs(k))));
9. s=strcat(['Escreva a taxa de variação da amplitude em porcentagem,
valor entre -100 e ', string(i-100), ': ']);
10. a=input(s);
11. H=strcat([N," volume variado em ",string(a-100),"%"]);
12. Y=wavwrite(((a+100)/100)*k,Fs,H);
13. sound([k,((a+100)/100)*k],Fs);
```

Descrição do algoritmo no Scilab para a variação do volume de áudio no formato .wav:

1. códigos para limpar o console de possíveis variáveis que possam estar carregadas, aumentar a pilha do Scilab ao máximo, e definir a resolução dos dados numéricos para contemplar todos os dígitos significativos.
2. código para pedir a entrada do áudio que terá o volume modificado.
3. código para carregar no Scilab os dados do áudio na matriz  $X$ , com a taxa de amostragem  $F_s$  e a resolução de cada amostra em  $bits$ , definidas na gravação (ou geração) do áudio.
4. código para carregar no Scilab a quantidade de linhas ( $nl$ ), e a quantidade de colunas ( $nc$ ) da matriz  $X$  (cada linha representa uma faixa do áudio).
5. código para concatenar a chamada da entrada da faixa com a linha da faixa de áudio a ser pedida como entrada.
6. código para carregar a entrada dos dados da(s) faixa(s) a ser(ere)m modificada(s).
7. código para carregar no console do Scilab os dados da(s) faixa(s) a ser(ere)m modificada(s).
8. código para determinar o fator em porcentagem que indica em quantos por cento pode-se aumentar a amplitude máxima do sinal para manter os dados do sinal de áudio com valores no intervalo  $[-1, 1]$ .
9. código para concatenar a chamada da entrada da taxa de variação da amplitude, com a taxa percentual máxima de acréscimo da amplitude do sinal dentro do intervalo  $[-1, 1]$ .
10. código para a entrada da taxa de variação do volume do sinal de áudio no intervalo  $[-1, 1]$ .
11. código para concatenar o nome do arquivo de entrada com a taxa de variação do volume (variação da amplitude).
12. código para gerar o áudio no formato .wav com o volume modificado e com a mesma taxa de amostragem  $F_s$  do arquivo de áudio original.
13. código para executar o áudio de entrada ( $X$ ) e na sequência o áudio de saída ( $Y$ ) com o volume modificado.

### 3.2.2 Aplicação do algoritmo para a variação do volume de áudio

Como exemplo de aplicação será usado como entrada um sinal determinístico em um arquivo de áudio .wav de uma senoide gerada pela função  $x(t) = 0,2\sin(2\pi 500t)$ , com  $t$  variando entre 0 s e 1 s, em uma frequência de 500 Hz, com amplitude máxima igual a 0,2 em uma escala linear com os dados numéricos do arquivo de áudio amostrados variando no intervalo  $[-1, 1]$ , coletados a uma taxa de amostragem de 44100 Hz e com resolução de 16 bits por amostra. A Figura 17 mostra os gráficos de onda das duas senoides, do som original e do som modificado pelo modelo, convertido em um gráfico contínuo pelo Audacity, sendo possível observar que, ao aumentar a amplitude em 200%, obtém-se como saída a senoide

$y(t) = 0,6\text{sen}(2\pi 500t)$ , no mesmo intervalo de tempo e com a mesma frequência, mas com amplitude máxima igual a 0,6 na escala linear, na mesma taxa de amostragem e com os mesmos *bits* de resolução por amostra.

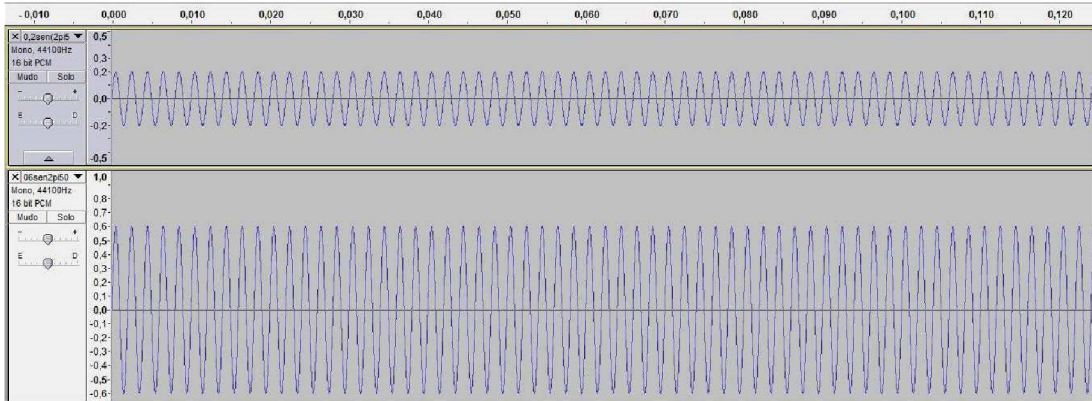


Figura 17: Variação da amplitude de onda gerada por senoide. Fonte: captura de tela do Audacity.

### 3.3 Envelopamento Dinâmico

O envelopamento dinâmico é uma forma de retratar o desenvolvimento dinâmico do som, do seu nascimento até a sua morte. Envelope dinâmico, também conhecido como envelope de amplitude, envoltório ou curva de envelope do som, representa esse desenvolvimento dinâmico através das curvas inferiores e superiores (Figura 18) que descrevem a variação da amplitude instantânea do som ao longo do tempo de áudio, em sua representação gráfica em forma de onda. O envoltório que delimita a representação gráfica do som, em geral, pode ser dividido em partes características: o ataque, é o momento inicial do som, decorrido entre o instante inicial e o momento onde o som atinge a amplitude máxima; o decaimento, que representa a perda de energia decorrente da grande energia liberada durante o ataque; a sustentação, é o momento onde ocorre a estabilização da intensidade, onde a amplitude tende a ser constante; e a extinção, quando o som tende ao fim [9, 40]. A duração desses momentos específicos do envelope dinâmico varia em cada tipo de som, o envelopamento é uma importante ferramenta durante o processo de modelagem matemática e computacional para simular e descrever a natureza do som, de forma geral, pode-se dizer que o envelopamento dinâmico representa a variação instantânea do volume do som.

Para representar essa variação de amplitude, natural ou não, utilizam-se a seguir funções matemáticas simples (nas Seções 3.3.1 e 3.3.2), e também funções mais flexíveis, obtidas por exemplo por função spline interpolante das amplitudes máximas locais de sons gerados naturalmente, como por exemplo, por um piano, uma flauta, uma marimba e por outros instrumentos musicais, essas funções serão



obtidas por “captura de envelope dinâmico” (Seção 3.3.3). O envelopamento dinâmico é uma das variantes a serem utilizadas no processo de modelagem.

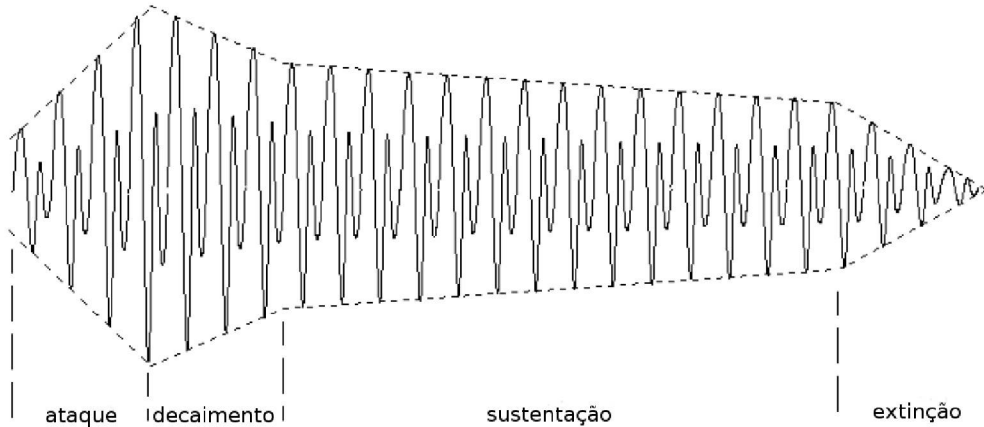


Figura 18: Representação das partes de um “Envelope Dinâmico” de uma onda sonora. Fonte: adaptado de [41].

### 3.3.1 Envelopamento do sinal de áudio por função afim

Um modelo para o envelopamento, na prática, consiste de uma variação não linear do volume, isto é, a variação da amplitude de cada amostra do sinal de forma desproporcional. Para facilitar a compreensão do processo de modelagem de sinal digital de áudio no domínio do tempo, será detalhado aqui nesta seção um envelopamento determinado por uma função afim e limitado superior e inferiormente por funções afins. Considere como sinal de entrada uma matriz  $X$  com  $n$  amostras de um arquivo de áudio (amplitudes instantâneas) dentro do intervalo  $[-1, 1]$  e com  $(t_f - t_0)$  segundos de duração. O objetivo deste modelo é variar por uma progressão aritmética as amplitudes desse sinal a partir de um instante inicial  $t_0$  até um instante final  $t_f$ , obtendo um sinal resultante, que, quando reconstruído, terá sua representação gráfica em forma de onda limitada por duas retas  $r$  e  $s$ , sendo a reta  $s$  simétrica a reta  $r$  em relação ao eixo das abscissas, por praticidade, a reta  $s$  será designada por  $-r$ . Na Figura 19 é possível ver uma representação gráfica de um sinal primitivo  $X$  na forma de senoide a ser envelopado por dois segmentos de reta  $r$  e  $s$ .



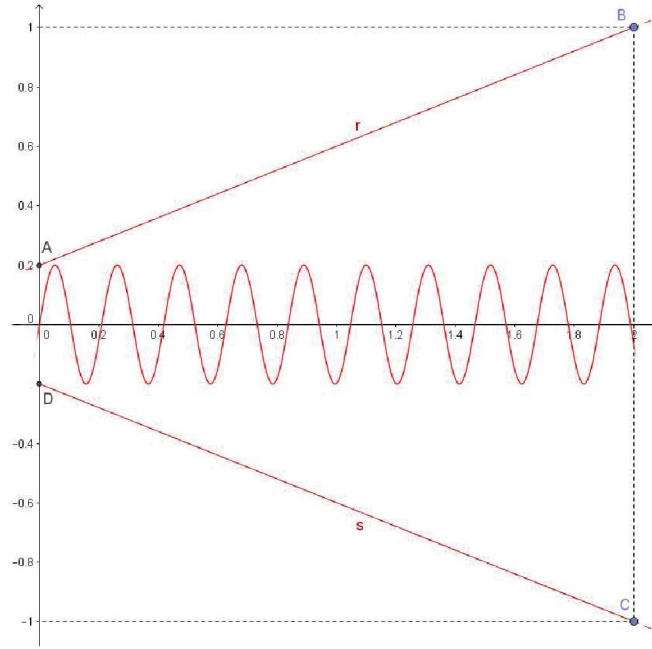


Figura 19: Representação de um sinal primitivo  $X$  (senoide) e dos segmentos de reta  $r$  e  $s$  que serão limitantes do sinal  $X$  após a aplicação do modelo para o envelopamento por função afim. Fonte: captura de tela do GeoGebra.

Uma função definida por  $f : \mathbb{R} \rightarrow \mathbb{R}$  chama-se afim quando existem constantes  $u, v$  que pertencem ao conjunto dos reais tais que  $f(x) = ux + v$  para todo  $x \in \mathbb{R}$ , sendo o seu gráfico representado por uma reta. O principal campo de atuação deste trabalho são os modelos matemáticos com domínio no tempo. A escala temporal, representada por  $t$ , tem como domínio o conjunto dos números reais não negativos ( $\mathbb{R}_+$ ), onde  $t$  representa cada instante de um sinal contínuo. Para ser possível discretizar um sinal contínuo, o mesmo deve ser limitado e de duração finita, desta forma, vamos obter uma função  $r$  a partir de uma função afim, mas com domínio apenas em  $[t_0, t_f] \subset \mathbb{R}_+$ , assim, o gráfico dessa função  $r$  será um segmento  $AB$  da reta  $r$ , com  $A = (t_0, r(t_0))$  e  $B = (t_f, r(t_f))$ . Considere uma função  $r : [t_0, t_f] \rightarrow \mathbb{R}_+$ , com  $t \in [t_0, t_f]$  e  $r(t) \in R \subset \mathbb{R}_+$ , associada aos pontos de uma progressão aritmética definida por  $\bar{r} : I \rightarrow \mathbb{R}_+$ , com  $I = \{i \in \mathbb{N}; 1 \leq i \leq n\}$  e  $\bar{r}(i)$  representando os elementos dessa progressão aritmética, tal que,  $\bar{r}(i) \in \bar{R} \subset R$ . Ou seja,  $R$  é um sinal contínuo e  $\bar{R}$  é um sinal discreto obtido por um determinado período de amostragem  $T = t_{i+1} - t_i$  desse sinal contínuo no intervalo de tempo  $[t_0, t_f]$ . Desta forma, pode-se associar o instante  $t$  com o tempo discreto  $iT$ , isto é,  $t = iT$  para algum  $T \leq t_f \in \mathbb{R}_+^*$ . Conhecendo a lei de formação da função contínua  $r$ , obtém-se um sinal contínuo e determinístico  $R$ , o que implica em  $r(t) = r(iT) \Leftrightarrow r[i] = \bar{r}(i)$ , ou seja, conhecendo a lei de formação da função contínua  $r$ , determina-se, pelo período de amostragem  $T$ , todos os  $n$  elementos da sequência  $\bar{R}$ , e conhecendo a sequência  $\bar{R}$ , determina-se a lei de

formação da função contínua  $r$  relacionada, ou seja, determina-se o sinal contínuo  $R$ . Supondo que  $r(t)$  representasse um sinal contínuo e aleatório, teria-se assim  $r(t) = r(iT) \Rightarrow r[i] = \bar{r}(i)$ , isto é, conhecendo o sinal contínuo  $R$ , determinaria-se pelo período de amostragem  $T$  todos os  $n$  elementos da sequência  $\bar{R}$ , mas conhecendo apenas a sequência  $\bar{R}$ , determinaria-se apenas uma aproximação do sinal contínuo  $R$  e essa aproximação poderia ser obtida por algum método de reconstrução de sinal contínuo, como por exemplo interpolação.

Mas sendo  $R$  um sinal determinístico, o que precisa ser feito, a princípio, é encontrar a função  $r$ , e depois deve-se discretizar  $R$  (conjunto imagem da função  $r$ ) para encontrar a sequência  $\bar{R}$ , e através do conjunto  $\bar{R}$ , conduzir as amplitudes instantâneas  $x(i)$  do conjunto discreto  $X$  para serem limitadas por uma função afim. Para que se tenha o controle do sinal a ser modelado, deve-se obter a amplitude máxima  $a$  do sinal  $X$ , o que pode ser feito no console do Scilab pelo código  $<a = \max(X)>$ , e define-se  $b \in \mathbb{R}_+$  como sendo  $b = r(t_f)$ , isto é,  $b$  é a imagem da função  $r$  no instante  $t_f$ . Para levar o sinal  $X$  a ser delimitado pelas retas  $r$  e  $-r$  é preciso obter uma função  $\bar{s}$ , discreta, que conduzirá cada elemento positivo  $x(i)$  da matriz  $X$  à aproximação de  $\bar{r}(i)$  e os elementos negativos à aproximação de  $-\bar{r}(i)$ . A função  $\bar{s}$  será chamada de “função envelopante”, isto é,  $\bar{s}$  é a função que leva cada  $|x(i)|$  a se aproximar de  $r[i]$ . A função  $r$  tem coeficiente linear dado por  $r(0)$  e dos pontos extremos do segmento de reta  $AB$  obtém-se a taxa de variação  $\frac{r(t_f)-r(t_0)}{t_f-t_0}$ , determinando-se, assim, a lei de formação da função  $r$  por

$$r(t) = \frac{r(t_f) - r(t_0)}{t_f - t_0} \cdot t + r(0), \quad (20)$$

onde neste algoritmo tem-se definido  $a = r(0)$ , para exercer o controle da amplitude instantânea máxima do sinal a ser envelopado. A seguir, vamos discretizar o conjunto  $R$ , ou seja, escrever  $r$  em função de  $i$ .

Dado  $j \in \mathbb{Z}_+$ , tal que,  $0 \leq j \leq n$  e  $t = jT$  para algum valor  $T \in \mathbb{R}_+^*$ . Pode-se, assim, reescrever a Equação 20 como

$$\begin{aligned} r(jT) &= \frac{r(nT) - r(T)}{nT} \cdot jT + r(0) \\ \Leftrightarrow r(jT) &= \frac{r(nT) - r(T)}{n} \cdot j + r(0T) \\ \Leftrightarrow r(jT) - r(0T) &= \frac{r(nT) - r(T)}{n} \cdot j, \end{aligned}$$

e sendo  $r(jT) - r(0T)$  e  $r(nT) - r(T)$  lineares, pode-se assim escrever<sup>15</sup>

$$T \cdot r(j) - T \cdot r(0) = \frac{T \cdot r(n) - T \cdot r(1)}{n} \cdot j \Leftrightarrow r(j) - r(0) = \frac{r(n) - r(1)}{n} \cdot j,$$

<sup>15</sup>Os valores  $r(n)$  e  $r(1)$  podem não ser definidos em  $[t_0, t_f]$ , mas são definidos na reta  $r$  que contém o segmento  $AB$ , pois uma função afim tem domínio em  $\mathbb{R}$ .

independente de  $T$ , e sendo por definição  $a = r(0)$ , escreve-se  $i = j$  se  $j \neq 0$ , obtendo assim o conjunto  $\bar{R}$  discretizado de  $R$  para qualquer período de amostragem  $T \leq t_f$ , dado por

$$\bar{r}(i) = \frac{\bar{r}(n) - \bar{r}(1)}{n} \cdot i + a. \quad (21)$$

O período de amostragem  $T$ , em geral, é um número muito pequeno<sup>16</sup>, sendo por exemplo, para uma taxa de amostragem de  $44100 \text{ Hz}$ ,  $t_0 \approx t_1$  e  $t_n \approx t_f$ . As funções afins limitantes aplicadas no modelo computacional a ser apresentado, em geral, têm taxas de variação não muito grandes, pois o sinal resultante precisa estar no intervalo  $[-1, 1]$ , podendo, assim, considerar as entradas dos valores  $a$  e  $b$  no algoritmo do modelo como sendo  $a = r(t_0) := r(t_1) = r(T) = \bar{r}(1)$  e  $b = r(t_f) := r(t_n) = r(nT) = \bar{r}(n)$ . Definindo  $a$  e  $b$  desta forma, pode-se reescrever a Equação 21 como

$$\bar{r}(i) \approx w(i) = \frac{b - a}{n} i + a,$$

que na linguagem do Scilab é escrita como

```
<r(i)=((b-max(X))/length(X))*i+max(X));>.
```

Tem-se  $\bar{r}(i) = w(i)$  se, e somente se,  $r[n] = r(t_f)$ . Para que a função  $r$  seja limitante, deve-se multiplicar cada elemento  $x(i) \in X$  pelo elemento de mesma ordem  $\bar{s}(i)$  de uma outra progressão aritmética  $\bar{s} : I \rightarrow \mathbb{R}_+$ , com  $\bar{s}(i) \in \bar{S}$ . Da mesma forma como foi feito com  $\bar{R}$ , vamos associar a progressão aritmética  $\bar{S}$  com a função  $s, s : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , com  $\bar{s}(i) = s[i] \in S$ . A taxa de variação da função  $\bar{s}$  será obtida multiplicando a taxa de variação do conjunto  $\{w(i) \in W; w(i) = \frac{b-a}{n}i + a\}$  por  $\frac{b}{a}$ , pois  $\frac{b}{a} \cdot a = b$ . Deseja-se, pelo algoritmo, que a amplitude instantânea do sinal  $X$  no instante inicial não tenha variação, isto é,  $s(0) \cdot x(0) = x(0)$  o que implica em  $s(0) = 1$ , sendo o coeficiente linear da função  $s$  igual a  $s(0)$ , chega-se na sequência

$$\bar{s}(i) = \frac{b}{a} \cdot \frac{b - a}{n} i + 1,$$

que, na linguagem do Scilab, é escrita como

```
<s(i)= (b/max(X))*((b-max(X))/length(X))*i+1>.
```

Obtém-se, assim, um conjunto  $Y_{1 \times n} = [y_1 \ y_2 \ y_3 \ \dots \ y_i \ \dots \ y_n]$  em função de  $i \in I$ , tal que,

$$y(i) = \bar{s}(i) \cdot x(i).$$

Sendo  $Y$  a matriz com os dados do áudio envelopado por função afim.

Como resultado, temos um áudio de saída gerado pela matriz  $Y$ , no formato .wav, com envelopamento por função afim. Apresenta-se a seguir o algoritmo no Scilab para envelopamento de um áudio por função afim.

---

<sup>16</sup>Para uma taxa de amostragem  $Fs = 44100 \text{ Hz}$ , tem-se  $T \approx 0,00002267573 \text{ s}$ .

### 3.3.1.1 Algoritmo no Scilab para envelopamento por função afim de um áudio no formato .wav

```

1. clear; stacksize('max'); format(16);
2. N=input("Escreva o nome do arquivo de audio wav:", "string");
3. b=input("Escreva a amplitude final do áudio(entre 0 e 1):");
4. [X,Fs,bits]=wavread(N);
5. C=strcat([N," com envelopamento por função afim"]);
6. k=[X(1,:)];
7. for i=1:length(k);
8.   s(i)=(b/max(k))*((b-max(k))/length(k))*i+1;
9.   Y(i)=s(i)*k(i);
10. end
11. wavwrite(Y,Fs,C);

```

Na Figura 20 pode-se ver uma representação gráfica, em sinal contínuo, de uma reta envelopante  $s$ , das retas  $r$  e  $-r$  limitantes superior e inferiormente, a curva  $y$ , representando o sinal de tempo discreto  $Y$ , a senoide  $x$ , representando o sinal primitivo  $X$ , além dos valores definidos na entrada do algoritmo,  $a = \max(X)$  e  $b$ , e sendo  $\bar{b}$  a amplitude máxima de  $Y$ .

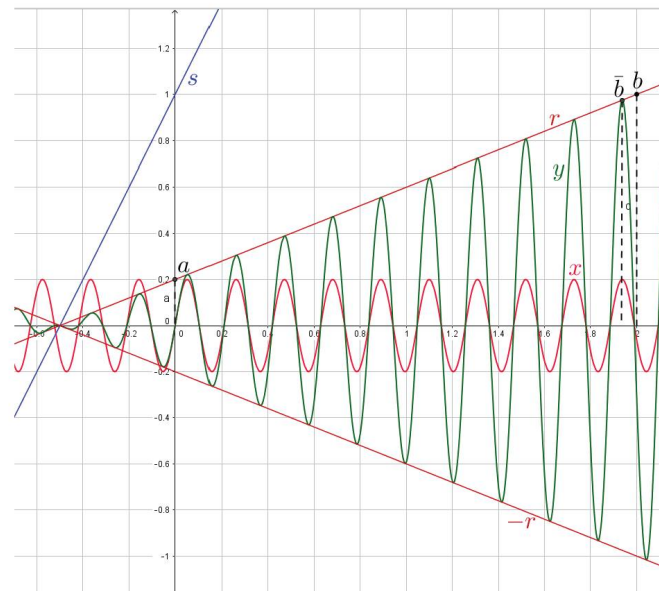


Figura 20: Senoide envelopada por função afim. Fonte: captura de tela do GeoGebra.



### 3.3.1.2 Resultado da aplicação do algoritmo para envelopamento de um áudio por função afim

Como exemplo, considere um áudio inicial com sinal  $X$ , na taxa de amostragem de  $44100 \text{ Hz}$ , com  $16 \text{ bits}$  de resolução de cada amostra e duração de  $0,5 \text{ s}$ , gerado por uma senoide obtida por

$$x(t) = 0,2 \sin(2\pi 100t), \quad (22)$$

onde o gráfico desta função pode ser visto na Figura 21.

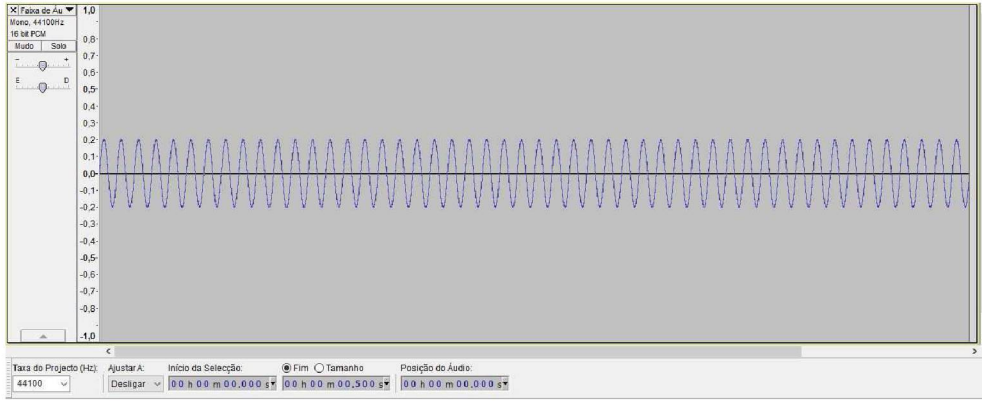


Figura 21: Senoide gerada no Audacity. Fonte: captura de tela do Audacity.

Aplicando o algoritmo do envelopamento por função afim para o sinal  $X$ , e definindo  $b = 1$ , tem-se como resultado um áudio gerado pelo sinal  $Y$ , envelopado por função afim, com amplitude variando de  $a$  até a amplitude final  $\bar{b} \leq b = 1$  (Figura 22). Para esta implementação, o envelope dinâmico têm as linhas superiores e inferiores geradas, por praticidade, por uma mesma função  $r$ , resultando nas linhas limitantes  $r$  e  $-r$ .

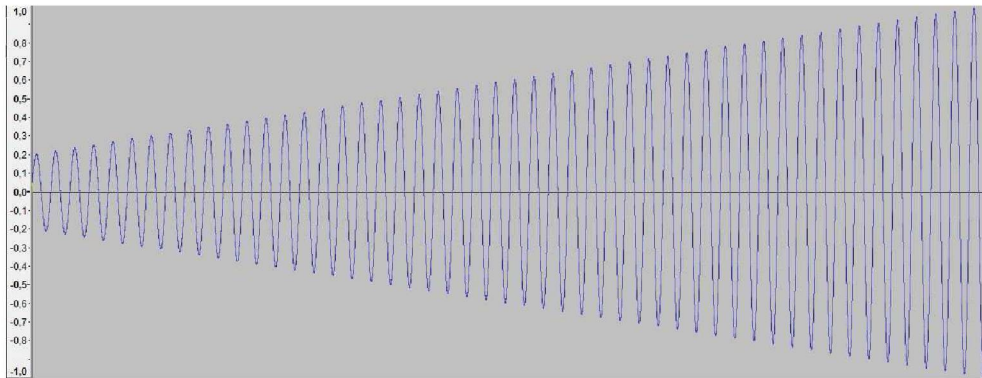


Figura 22: Senoide envelopada por função afim. Fonte: captura de tela do Audacity.

Durante todo o processo deve-se controlar  $\bar{b} = \max(Y)$  para que o mesmo esteja sempre no intervalo de áudio  $(0,1]$ , a condição inicial  $a = \max(X)$  serve para exercer esse controle, para não incorrer em perda indesejada de informação no sinal resultante  $Y$ , e o valor atribuído a  $b$  será determinante para que isso não ocorra.

### 3.3.2 Envelopamento do sinal de áudio por função exponencial

Outra possibilidade é fazer o envelopamento de um sinal de áudio por uma função exponencial, ou seja, variar exponencialmente o volume do áudio. Considere  $n$  amostras de som carregadas em uma matriz  $X$  com  $(t_f - t_0)$  segundos de duração. O objetivo é obter uma certa função exponencial  $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+^*$ , com  $t \in \mathbb{R}_+$  e sendo  $G = \{g(t) \in \mathbb{R}_+^*; g(t) = c \cdot e^{k \cdot t}\}$  associada a uma progressão geométrica  $\bar{g} : I \rightarrow [-1, 1]$ , com  $i \in I$  e  $\bar{G} = \{\bar{g}(i) \in [-1, 1]; \bar{g}(i) = c \cdot e^{k \cdot i}\}$ , tal que  $g[i] = \bar{g}(i)$  e sendo  $g$  e  $-g$  as curvas exponenciais limitantes superior e inferiormente do sinal resultante  $Y$ . A função exponencial gera um sinal determinístico, sendo possível obtê-la a partir de um sinal discreto associado. Considerando  $g(0) \approx g[1]$  e  $g[n] \approx g(t_f)$ , pelas mesmas razões relatadas na Seção 3.3.1 e para facilitar a resolução do sistema formado pelas Equações 23 e 24, mesmo não sendo definido  $\bar{g}(i)$  para  $i = 0$ , pode-se obter a função  $\bar{g}$  através das condições

$$\bar{g}(0) = \max(X) = a \quad (23)$$

e

$$\bar{g}(n) = b \quad (24)$$

onde tem-se que

$$\bar{g}(0) = c \cdot e^{k \cdot 0} = c \cdot e^0 = a,$$

o que implica em  $c = a$ . Da equação 24 tem-se

$$\bar{g}(n) = c \cdot e^{k \cdot n} = a \cdot e^{k \cdot n} = b,$$

sendo

$$\ln e^{k \cdot n} = \ln \frac{b}{a} \Rightarrow k = \frac{\ln \frac{b}{a}}{n}.$$

Resolvendo o sistema formado pelas Equações 23 e 24, e sendo  $g[i] = \bar{g}(i)$ , pode-se escrever a função

$$\bar{G} = \{\bar{g}(i) \in [-1, 1]; \bar{g}(i) = a \cdot e^{\frac{\ln \frac{b}{a}}{n} \cdot i}\},$$

que escrita em código para o Scilab torna-se

```
<g(i)=max(X)*exp(((log(b/max(X)))/length(X))*i);>.
```

Observa-se que, em código para o Scilab, a notação  $\ln$  é definida por  $\log$ . Para controlar a variação máxima de amplitude no envelope exponencial, deve-se também impor a variação  $\frac{b}{a}$  sobre cada  $\bar{g}(i)$ , obtendo  $S = \{s(i) \in \mathbb{R}_+; s(i) = \frac{b}{a} \cdot \bar{g}(i)\}$ , isto é,

$$s(i) = b \cdot e^{\frac{\ln \frac{b}{a}}{n} \cdot i},$$

sendo  $S$  a função discreta envelopante do sinal  $X$ , isto é, a função que levará cada elemento positivo  $x(i)$  da matriz  $X$  à aproximação de  $\bar{g}(i)$  e cada elemento negativo de  $X$  à aproximação de  $-\bar{g}(i)$ .

Pode-se, então, obter a matriz transformada

$$Y = \{y(i) \in [-1, 1]; y(i) = s(i) \cdot x(i)\},$$

onde  $Y$  é a matriz do sinal com envelope exponencial limitado pelas curvas exponenciais  $g$  e  $-g$ . A seguir, apresenta-se a implementação no código Scilab.

### 3.3.2.1 Algoritmo no Scilab para envelopamento por função exponencial de um áudio no formato .wav

```
1. clear; stacksize('max'); format(16);
2. N=input("Escreva o nome do arquivo de audio wav:", "string");
3. b=input("Escreva a amplitude final do áudio (entre 0 e 1):");
4. [X,Fs,bits]=wavread(N);
5. C=strcat([N," com amplificação exponencial"]);
6. k=[X(1,:)];
7. for i=1:length(k);
8.   s(i)=b*exp(((log(b/max(k)))/length(k))*i);
9.   Y(i)=s(i)*k(i);
10. end
11. wavwrite(Y,Fs,C);
```

### 3.3.2.2 Resultado da aplicação do algoritmo para envelopamento de um áudio por função exponencial

Como exemplo, considere um áudio inicial com sinal  $X$ , discretizado da Equação 22, na taxa de amostragem de  $44100 \text{ Hz}$ , com  $16 \text{ bits}$  de resolução de cada amostra e duração de  $0,5 \text{ s}$ . Após aplicar o algoritmo no sinal  $X$  definindo como entrada o valor  $b = 1$ , será obtido um áudio de saída gerado pela matriz  $Y$  no formato .wav com envelope exponencial, sendo o gráfico do sinal contínuo obtido do sinal  $Y$  limitado superior e inferiormente pelas curvas exponenciais  $g$  e  $-g$ , variando a amplitude máxima até  $\bar{b} \leq b$ , com  $\bar{b} \in (0, 1]$ . Obtém-se, neste exemplo, o áudio envelopado exponencialmente com amplitude máxima igual a  $\bar{b} \leq 1$ , onde, para  $t_f = 0,5 \text{ s}$ , tem-se  $n \approx Fs \cdot t_f$ , isto é,  $n = 44100 \cdot 0,5 = 22000$  amostras,

como pode ser observado no gráfico de sinal contínuo (Figura 23) construído a partir do sinal discreto  $Y$  com domínio em  $I$ .

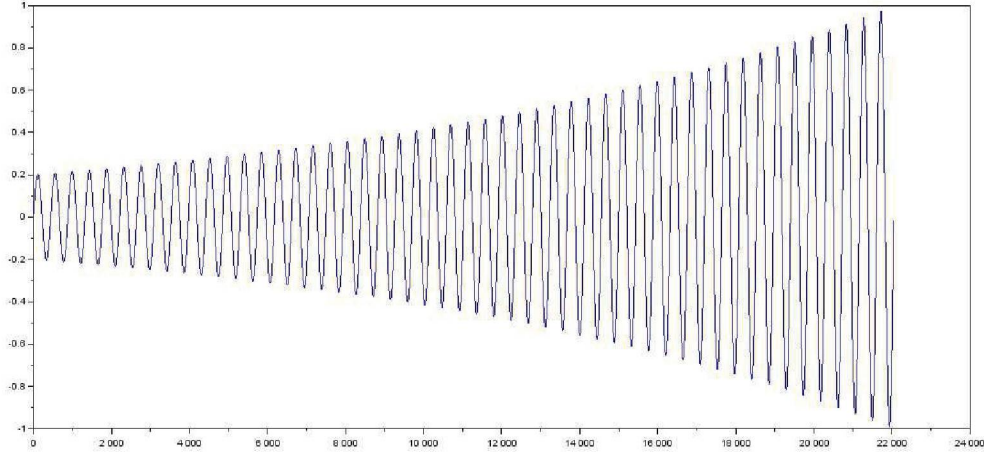


Figura 23: Senoide envelopada por função exponencialmente. Fonte: captura de tela do Scilab.

### 3.3.3 Envelopamento do sinal por funções mais flexíveis

Quando nos referimos a sons naturais, modelar um envelope dinâmico por funções afins, exponenciais ou por alguma outra função real tradicionalmente conhecida não é muito apropriado. As curvas limitantes que caracterizam o envelope dinâmico, tanto superior como inferior, são funções difíceis de serem moldadas por uma função específica (gerada por uma única equação, por exemplo). Uma opção é utilizar-se de “Funções Splines lineares Interpolantes<sup>17</sup>” para obter uma curva limitante, onde os nós, pontos de encontro de duas partes adjacentes da função spline, são os máximos locais do sinal  $X$ . Considere um sinal de áudio  $X$  gerado naturalmente (sinal aleatório), para obter os máximos locais do conjunto  $X$ , deve-se obter um certo subconjunto  $Y$ , com,  $y(i) \in Y$  e  $x(i) \in X$ , tal que,  $y(i) = x(i)$  se

$$x(i) \geq x(i-1) \quad (25)$$

e

$$x(i) \geq x(i+1), \quad (26)$$

além disso, deve-se obter os instantes  $t_i$  das  $i$ -ésimas amostras que satisfazem a condição dada pelas Desigualdades 25 e 26.

Uma função spline linear interpolante  $S_1(x)$  de  $y(x)$ , nos nós  $x_0, x_1, \dots, x_n$  é escrita em cada subintervalo  $[x_{i-1}, x_i]$ ,  $i = 1, 2, \dots, n$ , como

$$s_i(x) = y(x_{i-1}) \frac{x_i - x}{x - x_{i-1}} + y(x_i) \frac{x - x_{i-1}}{x_i - x_{i-1}}, \forall x \in [x_{i-1}, x_i]. \quad (27)$$

<sup>17</sup>Detalhes da demonstração da obtenção da função spline interpolante podem ser vistos em [43].



Por experiências realizadas, verificou-se, heurísticamente, que a utilização da spline linear tem um resultado melhor que uma spline cúbica, por exemplo, pois o erro de interpolação é visivelmente menor, o que pode ser observado quando se sobrepõem os gráficos do sinal original e dos envelopes superiores, justificando a opção pela spline linear interpolante.

O seguinte algoritmo é uma implementação no Scilab para obter as coordenadas dos pontos de máximos locais, e a função spline linear que interpola esses pontos. Como o conjunto de dados coletados  $X$  representa, em geral, pontos de uma curva bastante sinuosa em um curto intervalo de tempo, para obter um resultado melhor, aplica-se o algoritmo em quatro iterações, isto é, obtém-se um conjunto  $W$  com os máximos locais de  $X$ , na sequência  $W_1$ , com os máximos locais de  $W$ , e depois  $W_2$ , com os máximos locais de  $W_1$ , e finalmente  $W_3$ , com os máximos locais de  $W_2$ . A quantidade de iterações foi decidida de forma subjetiva<sup>18</sup>, pela análise gráfica dos resultados.

### 3.3.3.1 Implementação no Scilab para a captura de envelope dinâmico

```

1. clear; stacksize('max'); format(16);
2. z=input("Escreva o nome do arquivo de audio envelopante wav:","string");
3. [k,fs,bits]=wavread(z);
4. X=[k(1,:),0,0];
5. for i=1:length(X)-2
6.     if X(i)<=X(i+1) & X(i+1)>=X(i+2) then
7.         a(i+1)=X(i+1);
8.     end
9. end
10. b=[[1:length(a)]',a];
11. for j=1:length(a)
12.     if b(j,2)>0 then
13.         c(j,:)=b(j,:);
14.     end
15. end
16. u=setdiff(c(:,1),0);
17. m=0
18. for n=1:length(a)
19.     if a(n)>0 then
20.         m=m+1;
21.         v(m)=a(n);
22.         d(m)=n;
23.     end
24. end

```

---

<sup>18</sup>Com 5 iterações ou mais o erro de interpolação aumenta consideravelmente, e esse erro é visível quando se sobrepõe o gráfico do sinal de entrada  $X$  por  $W_j$ , com  $1 \leq j \leq 5$ .

```

25. t=1:length(X);
26. w=interp1n([u';v'],t');
27. for i1=1:length(w)-2
28.     if w(i1)<=w(i1+1) & w(i1+1)>=w(i1+2) then
29.         a1(i1+1)=w(i1+1);
30.     end
31. end
32. b1=[[1:length(a1)]',a1];
33. for j1=1:length(a1)
34.     if b1(j1,2)>0 then
35.         c1(j1,:)=b1(j1,:);
36.     end
37. end
38. u1=setdiff(c1(:,1),0);
39. m1=0
40. for n1=1:length(a1)
41.     if a1(n1)>0 then
42.         m1=m1+1;
43.         v1(m1)=a1(n1);
44.         d1(m1)=n1;
45.     end
46. end
47. w1=interp1n([u1';v1'],t');
48. for i2=1:length(w1)-2
49.     if w1(i2)<=w1(i2+1) & w1(i2+1)>=w1(i2+2) then
50.         a2(i2+1)=w1(i2+1);
51.     end
52. end
53. b2=[[1:length(a2)]',a2];
54. for j2=1:length(a2)
55.     if b2(j2,2)>0 then
56.         c2(j2,:)=b2(j2,:);
57.     end
58. end
59. u2=setdiff(c2(:,1),0);
60. m2=0
61. for n2=1:length(a2)
62.     if a2(n2)>0 then
63.         m2=m2+1;
64.         v2(m2)=a2(n2);
65.         d2(m2)=n2;
66.     end
67. end

```

```

68. w2=interp1n([u2';v2'],t');
69. for i3=1:length(w2)-2
70.     if w2(i3)<=w2(i3+1) & w2(i3+1)>=w2(i3+2) then
71.         a3(i3+1)=w2(i3+1);
72.     end
73. end
74. b3=[1:length(a3)]',a3];
75. for j3=1:length(a3)
76.     if b3(j3,2)>0 then
77.         c3(j3,:)=b3(j3,:);
78.     end
79. end
80. u3=setdiff(c3(:,1),0);
81. m3=0
82. for n3=1:length(a3)
83.     if a3(n3)>0 then
84.         m3=m3+1;
85.         v3(m3)=a3(n3);
86.         d3(m3)=n3;
87.     end
88. end
89. w3=interp1n([u3';v3'],t');
90. subplot(311);plot2d(t',X',rect=[min(t),min(y),max(t),max(y)]);
91. subplot(312);plot2d(t',w1,rect=[min(t),min(y),max(t),max(y)]);
92. subplot(313);plot2d(t',w3,rect=[min(t),min(y),max(t),max(y)]);

```

### 3.3.3.2 Resultado da aplicação do algoritmo para a captura de envelope dinâmico

Como exemplo, será usado um sinal de entrada  $X$  obtido da gravação digital do som da terceira oitava<sup>19</sup> da nota musical Dó de um piano ( $C_3$ ), com frequência aproximada de  $130,8 \text{ Hz}$  e período de  $0,007645 \text{ s}$  aproximadamente.

Como resultado, tem-se os gráficos na Figura 24, onde o primeiro gráfico representa o sinal contínuo construído a partir do sinal discreto  $X$  com domínio em  $I$ , o segundo gráfico representa o sinal contínuo construído a partir de  $W$ , obtido por spline linear interpolante dos máximos locais do sinal  $X$ , e  $X$  com domínio em  $I$ , e o terceiro gráfico é a representação contínua (linha) da iteração  $W_3$  do algoritmo, que representa uma aproximação do envelope superior do sinal  $X$ . Observa-se no gráfico algumas partes do envelope com pontos discrepantes

---

<sup>19</sup>Em [9],  $C_3$  corresponde a nota musical com frequência de  $130,8 \text{ Hz}$ , porém, em alguns instrumentos musicais eletrônicos e também em outras referências, como por exemplo em [46], a frequência de  $130,8 \text{ Hz}$  está associada a nota  $C_2$ .

ao comportamento geral da curva. Para corrigir problemas deste tipo, pode-se aplicar mais iterações apenas nos subintervalos onde os erros de interpolação se acentuaram, refinando o resultado e obtendo um envelope mais apropriado.

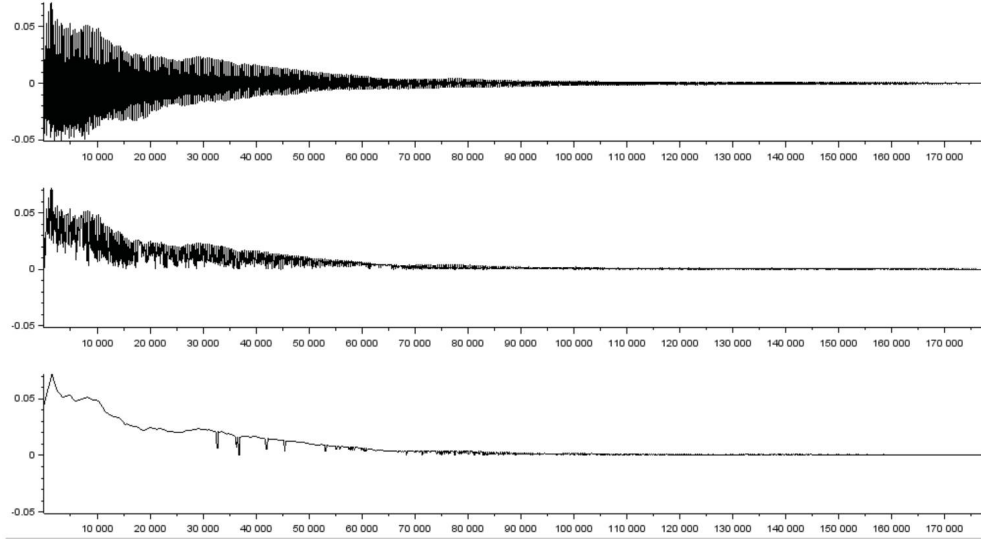


Figura 24: Captura do envelope superior do sinal de áudio de um piano. Fonte: captura de tela do Scilab.

Após adquirir, por função spline linear, o conjunto de pontos  $W_j$ , que é um conjunto discreto, pode-se obter um conjunto contínuo  $M \supset W_j$  por função spline cúbica interpolante<sup>20</sup> nos nós dados pelos pontos do conjunto  $W_j$ . E, posteriormente, um conjunto discreto  $Q \subset M$ , com a mesma quantidade de amostras e a mesma taxa de amostragem de uma sinal  $\bar{X}$  para o qual pretende-se aplicar o envelopamento. De forma análoga procedida na Implementação 3.3.2.1, substituindo o conjunto  $\bar{G}$ , com  $\bar{g}(i) \in \bar{G}$ , desse algoritmo pelo conjunto  $Q$ , com  $q(i) \in Q$ , e sendo  $a$  a amplitude máxima de  $\bar{X}$  e  $b$  a variável que determinará a amplitude máxima do sinal resultante  $Y$ , tomando o devido cuidado ao escolher o valor  $b$ , para as amplitudes instantâneas não extrapolarem o intervalo  $[-1,1]$ , ocasionando em perda de informação. Obtém-se, assim, um novo sinal  $Y$  envelopado pelo envelope capturado do sinal original  $X$ , sendo necessário, dependendo da amplitude máxima do sinal  $\bar{X}$ , ajustar o volume do sinal resultante  $Y$ , para ser possível ouvir o áudio resultante.

### 3.4 Efeitos de áudio para reverberação, delay e eco

Outra possibilidade é modelar repetições do sinal de áudio com atrasos definidos, as quais simulam a sensação de um áudio gerado em um ambiente onde ocorrem

<sup>20</sup>Detalhes da obtenção de uma função spline cúbica interpolante na Seção 3.5.



reflexões das ondas sonoras. Essas reflexões, quando são captadas em intervalos definidos de tempo, contribuem para a sensação de reverberação, delay ou eco do áudio.

O fenômeno da reverberação (reverb) acontece, por exemplo, quando após cessado um som gerado em uma sala, continuamos ainda a escutá-lo por um determinado tempo. O efeito reverb é o resultado de sucessivas reflexões do som, que se somam com pequenos atrasos para provocar a sensação percebida. Além dos atrasos, as ondas sonoras sofrem um decaimento de amplitude ao longo do tempo, devido, principalmente, à absorção acústica pelas superfícies refletoras e às sucessivas reflexões, que aumentam a distância de deslocamento das ondas até serem captadas, ocasionando uma maior absorção das intensidades sonoras pelo ar [5].

Quando ocorrem as primeiras reflexões das ondas sonoras, e essas são captadas em um atraso compreendido entre 10 *ms* e 100 *ms* aproximadamente, é gerado um reforço sonoro natural, provocado pelo pequeno intervalo de tempo entre a captação direta do som e a captação das primeiras reflexões. As reflexões sucessivas que são consequentes das primeiras reflexões caracterizam o que se define como reverberação, e esse fenômeno pode ser sintetizado por modelagem de onda sonora. Numa acústica real ocorrem também atenuações das amplitudes das frequências parciais mais altas das ondas refletidas, que variam proporcionalmente à quantidade de repetições e pelo tipo de material refletor.

Alguns parâmetros podem ser usados para descrever matematicamente esse fenômeno: o pré-atraso, que é o tempo decorrido entre a captação direta do som e as suas primeiras reflexões; a densidade de reverberação, que determina a quantidade de superfícies difusoras do ambiente; o tamanho da sala ambiente, que interfere no intervalo dos sucessivos atrasos; e a difusão, relacionada às propriedades do revestimento do ambiente, que implica em sucessivas atenuações da amplitude devido a absorção da energia da onda a cada reflexão. Estes parâmetros se refletem na estrutura do “efeito reverb”, que se divide basicamente em três regiões: o som direto no menor intervalo de tempo, as primeiras reflexões e a região de reverberação, que são as sucessivas reflexões decorrentes das primeiras reflexões, como podemos observar na Figura 25.

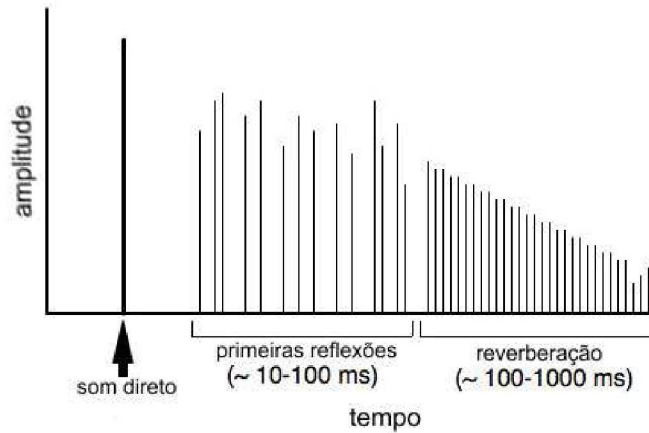


Figura 25: Representação do efeito reverb: o som direto (áudio original), sem interferência da acústica do ambiente, é seguido por repetições posteriores, com amplitudes atenuadas. Fonte: adaptado de [44].

O efeito reverb é um recurso antigo em produção de áudio, presente em vários tipos de mesas de som, e é um processo onde, em geral, não é necessário digitalizar o áudio. O efeito reverb analógico é simulado quando o sinal de áudio é armazenado em um buffer eletrônico, para então ser reenviado para uma saída de áudio, mas com um pequeno atraso de tempo, resultando em um áudio com efeito de reverberação. Nesse processo não é necessário um recurso computacional, porém, é um procedimento de baixa precisão, pois não se tem o controle exato dos atrasos das repetições. No modelo apresentado, o atraso sonoro é gerado por um modelo matemático baseado em manipulação matricial de dados numéricos de um arquivo de áudio digital, e é um modelo bastante preciso.

O modelo que é apresentado aqui destina-se a gerar uma matriz de dados de intensidades sonoras, onde em cada linha esteja representada a faixa de áudio direto, e os áudios com amplitudes atenuadas progressivamente e em intervalos de atrasos definidos para as faixas subsequentes. O objetivo prático é simular o efeito reverb em um áudio captado a uma taxa de amostragem  $Fs$ , com resolução de  $b$  bits, e  $n$  amostras de intensidade sonora equiespaçadas num intervalo de tempo  $t$ .

Considere então como entrada um áudio monoaural com os dados armazenados em uma matriz  $X_{1 \times n} = [x_1 \ x_2 \ x_3 \ \dots \ x_i \ \dots \ x_n]$ . Para sintetizar o efeito de reverberação, devemos definir os atrasos  $r_j$ , com  $r_j \leq r_{j+1}$ , em milissegundos, em um vetor  $R_{1 \times m} = [r_1 \ r_2 \ r_3 \ \dots \ r_j \ \dots \ r_m]$ . Desta forma,  $L = m + 1$  representará a quantidade de faixas de áudio a ser usada para simular o reverb, a faixa primitiva mais  $m$  faixas com atrasos  $r_j$ ,  $1 \leq j \leq m$ . Além disso, define-se a taxa de decréscimo  $p_j$  de cada amplitude em porcentagem, para serem aplicadas nos áudios com atraso, pelo vetor  $P_{1 \times m} = [p_1 \ p_2 \ p_3 \ \dots \ p_j \ \dots \ p_m]$ .

Para modelar o atraso sonoro é necessário introduzir em cada faixa de áudio

adicional um período de silêncio, que representa o tempo de atraso na execução do áudio. Este silêncio é modelado por vetores nulos que serão introduzidos em cada faixa de áudio. Vamos representar os períodos de atrasos, para a taxa de amostragem  $F_s$ , em cada faixa de áudio, pelo vetor nulo  $0(r_j)$ , onde a ordem de  $0(r_j)$  deve ser a parte inteira do número  $(r_j \cdot F_s)/1000$ .

O tempo de reverberação  $TR$ , dado em segundos, é determinado levando-se em consideração os coeficientes de absorção de cada material refletor  $\alpha_j$ , as áreas das superfícies da sala revestidas com cada material  $A_j$ , em  $m^2$ , em uma sala com volume  $V$ , em  $m^3$ , e é dado pela fórmula  $TR \approx 0.163 \cdot V / \sum \alpha_j \cdot A_j$ , conhecida como fórmula de Sabine [5]. Leva-se em consideração também a absorção da intensidade sonora do meio onde a onda se propaga, que neste caso, é o ar. A variável  $TR$  representa o tempo necessário para a amplitude da onda diminuir 60  $dB$ , ou seja,  $TR$  representa a diferença aproximada do tempo de duração do som reverberado com o som sem o efeito reverb, desta forma, pode-se definir o atraso máximo como  $r_m \approx TR$ .

Para percepção do efeito reverb é necessária a execução simultânea da faixa de áudio original e de  $m$  alterações da mesma pela introdução dos  $m$  atrasos  $r_j$  e das  $m$  atenuações de amplitude de  $p_j$  por cento. Portanto, deve-se gerar uma matriz de saída  $Y$ , que será a informação numérica de um novo arquivo de áudio, e cujas linhas são a faixa original e as faixas modificadas, as quais necessitam ter o mesmo tempo de duração. Por isso, todos os vetores nulos  $0(r_j)$ ,  $1 \leq j \leq m$ , são introduzidos na faixa original e em cada uma das faixas modificadas, porém em posições distintas: na faixa original se localizam nas últimas colunas e nas faixas modificadas nas primeiras e nas últimas colunas. Desta forma, a matriz de saída  $Y$  deverá ser de ordem  $L \times C$ , onde  $C = n + [\sum (r_j \cdot F_s)/1000]$ , para  $1 \leq j \leq m$ , ou seja, a soma das  $n$  amostras do áudio primitivo mais a parte inteira do somatório  $\sum (r_j \cdot F_s)/1000$ . A primeira linha da matriz  $Y$  deverá ser o vetor  $X$  acrescido de entradas nulas (zeros) nas últimas colunas. Isto é,

$$Y_{1 \times C} = [X_{1 \times n} \mid 0(r_1) \mid 0(r_2) \mid 0(r_3) \mid \dots \mid 0(r_j) \mid \dots \mid 0(r_m)].$$

A linha  $j$  de  $Y$  tem como entradas primeiro o vetor  $0(r_j)$  que carrega o atraso de tempo  $r_j$  em segundos, seguido do vetor  $(1 - p_j/100) \cdot X_{1 \times n}$ , que representa as  $n$  amostras  $x_i$  atenuadas em  $p_j$  por cento, e por último os demais vetores  $0(r_p)$ , com  $p \neq j$  a saber,

$$Y_{j \times C} = [0(r_j) \mid (1 - p_j/100) \cdot X_{1 \times n} \mid 0(r_1) \mid 0(r_2) \mid 0(r_3) \mid \dots \mid 0(r_{j-1}) \mid 0(r_{j+1}) \mid \dots \mid 0(r_m)],$$

onde  $1 \leq j \leq m$ . A condição  $r_j \leq r_{j+1}$  é imposta para que se tenha  $r_m$  como sendo o maior atraso, podendo-se assim, defini-lo a partir do tempo de reverberação  $TR$ .

A mesma relação de ordem não se aplica para os elementos  $p_j$ , pois, sendo atenuadores de amplitudes, sofrem influência do tipo de material refletor, ocasionando em uma onda com um atraso maior e com amplitude maior que outra com atraso menor.



Portanto, o áudio reverberado será modelado pela matriz

$$Y_{L \times C} = \begin{bmatrix} X_{1 \times n} | 0(r_1) | \dots | 0(r_m) \\ 0(r_1) | (1 - p_1/100) \cdot X_{1 \times n} | 0(r_2) | 0(r_3) | \dots | 0(r_m) \\ 0(r_2) | (1 - p_2/100) \cdot X_{1 \times n} | 0(r_1) | 0(r_3) | \dots | 0(r_m) \\ \vdots \\ 0(r_j) | (1 - p_j/100) \cdot X_{1 \times n} | 0(r_1) | \dots | 0(r_{j-1}) | 0(r_{j+1}) | \dots | 0(r_m) \\ \vdots \\ 0(r_m) | (1 - p_m/100) \cdot X_{1 \times n} | 0(r_1) | \dots | 0(r_{m-1}) \end{bmatrix}. \quad (28)$$

Neste modelo apresentado não se considerou a filtragem das frequências altas, gerada naturalmente durante as reflexões de ondas.

Na seção seguinte será apresentada uma implementação no Scilab para gerar um delay (atraso) com amplitude atenuada para simular o efeito reverb. O arquivo primitivo de áudio deverá ser, neste caso, estereofônico, e para esta implementação utiliza-se, para cada canal, apenas uma repetição, com atraso definido em milissegundos e com a taxa de decréscimo das amplitudes em porcentagem.

#### 3.4.1 Implementação no Scilab para modelagem do efeito reverb

```
1. clear; stacksize('max'); format(18);
2. m=input("Escreva o nome do arquivo de audio wav estereo:
","string");
3. r=input("Escreva o atraso do som do canal esquerdo em
milissegundos:");
4. q=input("Escreva o atraso do som do canal direito em
milissegundos:");
5. p1=input("Escreva a taxa de decréscimo da amplitude do
áudio com atraso do canal esquerdo (valor em porcentagem):");
6. p2=input("Escreva a taxa de decréscimo da amplitude do
áudio com atraso do canal direito (valor em porcentagem):");
7. [X,Fs,bits]=wavread(m);
8. E=X(1,:);
9. D=X(2,:);
10. [nl,nc]=size(E);
11. [l1,c1]=size(0:((r/1000)/((r/1000)*Fs-1)):((r/1000)));
12. [l2,c2]=size(0:((q/1000)/((q/1000)*Fs-1)):((q/1000)));
13. n=strcat([m," com delay de ",string(r)," ms e ",
string(q)," ms para cada faixa"]);
14. Y=[E,zeros(nl,c1),zeros(nl,c2);D,zeros(nl,c1),
zeros(nl,c2);zeros(nl,c1),(1-p1/100)*E,zeros(nl,c2);
zeros(nl,c2),(1-p2/100)*D,zeros(nl,c1)];
15. wavwrite(Y,Fs,n);
```



```

16. playsnd([X,[Y(1,:);Y(3,:)], [Y(2,:);Y(4,:)]],Fs);
17. t=1/Fs:1/Fs:(length([E,zeros(nl,c1),zeros(nl,c2)])/Fs);
18. subplot(221); xtitle('Áudio no canal esquerdo');
plot2d(t,[E,zeros(nl,c1),zeros(nl,c2)],
rect=[0,min([E,D]),length([E,zeros(nl,c1),zeros(nl,c2)])
/Fs,max([E,D]])];xlabel("segundos");ylabel("amplitude");
subplot(222); xtitle('Áudio no canal direito'); plot2d(t,
[D,zeros(nl,c1),zeros(nl,c2)],rect=[0,min([E,D]),length(
[D,zeros(nl,c1),zeros(nl,c2)])/Fs,max([E,D]])];xlabel(
"segundos");ylabel("amplitude");
subplot(223); xtitle(strcat(['Faixa do canal esquerdo
com atraso de ',string(r),' ms e com amplitude atenuada
em ',string(p1),' %'])); plot2d(t,[zeros(nl,c1),(1-p1/100)
*E,zeros(nl,c2)],rect=[0,min([E,D]),length([E,zeros(nl,c1),
zeros(nl,c2)])/Fs,max([E,D]])];xlabel("segundos");ylabel("
amplitude");subplot(224); xtitle(strcat(['Faixa do canal
direito com atraso de ',string(q),' ms e com amplitude
atenuada em ',string(p2),' %']));plot2d(t,[zeros(nl,c2),
(1-p2/100)*D,zeros(nl,c1)],rect=[0,min([E,D]),length([E,
zeros(nl,c1),zeros(nl,c2)])/Fs,max([E,D]])];xlabel("
segundos");ylabel("amplitude");

```

### 3.4.2 Descrição do algoritmo de modelagem do efeito reverb

1. códigos para limpar o console de possíveis variáveis que possam estar carregadas, aumentar a pilha do Scilab ao máximo e definir a resolução dos dados numéricos para contemplar todos os dígitos significativos;
2. código para pedir a entrada de um arquivo de áudio com formato .wav;
3. código para pedir a entrada do tempo de atraso do áudio da faixa do canal esquerdo em milissegundos (ms);
4. código para pedir a entrada do tempo de atraso do áudio da faixa do canal direito em milissegundos (ms);
5. código para pedir a taxa de decréscimo da intensidade sonora, para o áudio da faixa do canal esquerdo com atraso, em porcentagem;
6. tem-se o código para pedir a taxa de decréscimo da intensidade sonora, para o áudio da faixa do canal direito com atraso, em porcentagem;
7. código para carregar no console do Scilab os dados do áudio na matriz X, os argumentos Fs (taxa de amostragem) e bits (resolução de cada amostra);
8. código para selecionar em uma nova matriz a linha 1 da matriz X que representa a faixa de áudio do canal esquerdo;
9. código para selecionar em uma nova matriz a linha 2 da matriz X que representa a faixa de áudio do canal direito;
10. código para determinar o número de linhas (nl) e de colunas (nc) da matriz

- k1 (todas as faixas de entrada têm a mesma quantidade de elementos);
11. código para determinar o número de linhas (l1) e de colunas (c1) da matriz zero que condicionará ao atraso pretendido à faixa 1;
  12. código para determinar o número de linhas (l2) e de colunas (c2) da matriz zero que condicionará ao atraso pretendido à faixa 2;
  13. código para concatenar o nome do arquivo de entrada com o efeito gerado;
  14. código para gerar a matriz do áudio reverberado, em dois canais, o primeiro com o áudio original e o segundo com o atraso de amplitude atenuada;
  15. código de como gerar o áudio da matriz Y, com a mesma taxa de amostragem da entrada (Fs), e com nome de saída concatenado em n;
  16. código para executar o áudio de entrada e os de saída com delay, uma faixa de áudio por vez para avaliação;
  17. código para determinar a matriz t com os instantes do áudio com delay;
  18. código para gerar o gráfico das faixas de áudio original junto com os áudios modificados.

### 3.4.3 Aplicação do algoritmo para modelagem do efeito reverb

O modelo matemático apresentado considera como entrada um arquivo de áudio digital com apenas uma faixa de áudio. Entretanto, a partir deste modelo básico, com algumas pequenas modificações podemos reformulá-lo para receber mais de uma faixa do áudio primitivo, mantendo a estrutura básica do algoritmo, substituindo as primeiras faixas de áudio atenuadas com atrasos por faixas adicionais do áudio original, e obtendo os demais áudios atenuados com atraso a partir de cada faixa de áudio primitivo igualmente distribuído. Dessa forma, pode-se trabalhar com áudio estereofônico, por exemplo, com canais independentes, tendo como saída, as duas primeiras faixas de áudio original, representadas pelas duas primeiras linhas da matriz Y, em cada um dos canais, esquerdo e direito, respectivamente. As demais faixas, representadas pelas outras linhas da matriz Y, terão saída monoaural, isto é, o áudio executado sairá igualmente nos dois canais.

Na simulação mostrada na Figura 26 utiliza-se como sinal original um arquivo de áudio estéreo, gravado em duas faixas, uma em cada canal. Cada faixa foi obtida por microfones distintos com profundidade de captação diferente, um dinâmico e outro condensador, gravados no mesmo instante, no software livre de manipulação de áudio Audacity. A taxa de amostragem definida na captação foi de 44100 Hz, com resolução de 16 bits por amostra, no formato de áudio .wav, com a gravação “O efeito delay é obtido pela soma do sinal original com o sinal atrasado”. Pode-se observar que os gráficos superiores (direito e esquerdo) mostram respectivamente as amplitudes dos dois canais esquerdo e direito da gravação estéreo original, enquanto que os gráficos inferiores, mostram que cada um dos sinais originais sofreram atrasos de 60 ms e 80 ms, respectivamente, bem como atenuações de amplitude de 30% e 50%, respectivamente. A execução simultânea de todas estas faixas de áudio simula o efeito sonoro reverb.

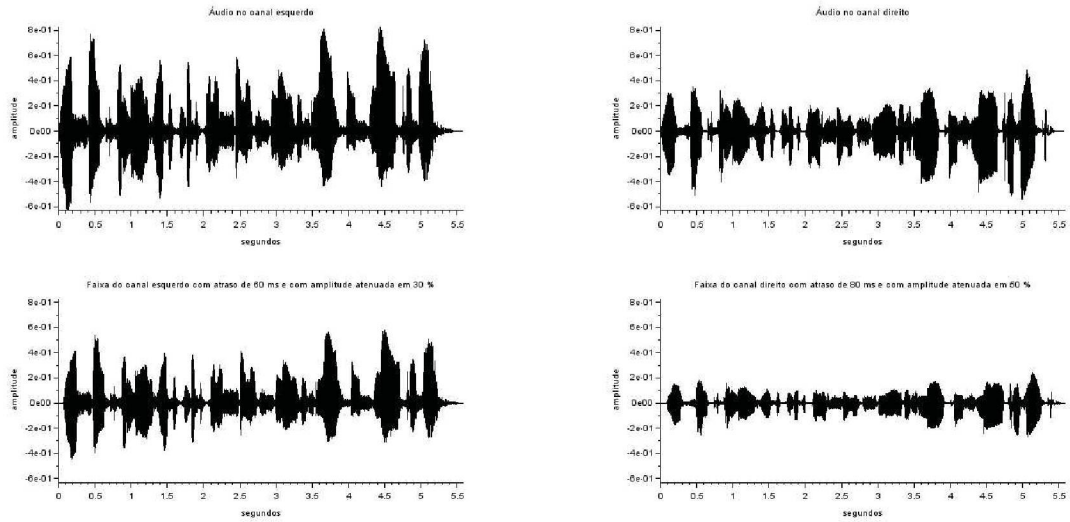


Figura 26: Simulação de reverb para o áudio estéreo da frase “O efeito delay é obtido pela soma do sinal original com o sinal atrasado”. Fonte: captura de tela do Scilab.

#### 3.4.4 Considerações sobre o modelo matemático do efeito reverb

O modelo matemático apresentado além de simular o efeito sonoro do fenômeno natural reverb, também pode ser aplicado para simular o fenômeno eco. O efeito eco é obtido quando o atraso no áudio excede um determinado limite em *ms*, que faz com que se perca a sensação natural de reforço do áudio, típico do reverb, sendo possível distinguir no eco, com clareza, cada repetição [5, 2]. Tanto o eco como o reverb apresentam o decaimento de amplitude nas repetições, tendo as mesmas características, e, portanto, podem ser modelados por um mesmo modelo matemático. O delay, ao contrário do reverb e do eco, não se pode classificar como um fenômeno natural, pois uma onda sonora tem que se propagar em um meio físico mínimo, entre a fonte e o captador de som, para se caracterizar como onda sonora. E no percurso da onda no meio material, por menor que seja, ocorrerá alguma perda de intensidade. Dessa forma, na prática, não temos como gerar um delay natural. Portanto, definimos delay apenas como efeito, e não como fenômeno. O efeito delay pode ser facilmente implementado pelo algoritmo descrito pela Equação 28, basta definir  $p_j = 0$ , para todo  $j$ , sendo  $p_j$  a taxa de decréscimo de amplitude aplicada em cada faixa de áudio com atraso, obtendo assim apenas as repetições, sem a atenuação de amplitude, isto é, apenas o delay.

### 3.5 Modelagem de ondas sonoras aplicada à música

A nota musical é um tom, e um tom é formado por uma frequência ou pela combinação de uma pequena quantidade de frequências. O tom de uma nota musical



é um som que agrada os nossos ouvidos. Em música, tom é definido como uma distância de dois sustenidos (ou de dois bemóis), e um semitom é uma distância de um sustenido (ou de um bemol). As escalas musicais são sequências ordenadas de notas. Em geral, as escalas musicais se dividem em “escalas maiores” e “escalas menores”, existindo para cada tipo várias escalas. Por exemplo, repetindo o ciclo, Dó, Ré, Mi, Fá, Sol, Lá, Si, Dó ..., na sequência tom, tom, semitom, tom, tom, tom, semitom ..., obtemos uma escala maior de Dó. Para essa mesma sequência de tons e semitons podemos escrever uma escala maior para outras notas, por exemplo, a sequência, Lá, Si, Dó<sup>#</sup>, Ré, Mi, Fá<sup>#</sup>, Sol<sup>#</sup>, Lá ..., e repetindo o ciclo, obtemos uma escala maior em Lá. Pode-se assim, construir uma escala maior de todas as 12 notas conhecidas, na Figura 27 temos a escala maior das 7 notas básicas [45].

<i>Intervalos da escala</i>	<i>Tom</i>	<i>Tom</i>	<i>Semi-Tom</i>	<i>Tom</i>	<i>Tom</i>	<i>Tom</i>	<i>Semi-Tom</i>	
<b>Escala Dó maior</b>	C	D	E	F	G	A	B	C
<b>Escala Ré maior</b>	D	E	F <sup>#</sup>	G	A	B	C <sup>#</sup>	D
<b>Escala Mi maior</b>	E	F <sup>#</sup>	G <sup>#</sup>	A	B	C <sup>#</sup>	D <sup>#</sup>	E
<b>Escala Fá maior</b>	F	G	A	B <sup>b</sup>	C	D	E	F
<b>Escala Sol maior</b>	G	A	B	C	D	E	F <sup>#</sup>	G
<b>Escala Lá maior</b>	A	B	C <sup>#</sup>	D	E	F <sup>#</sup>	G <sup>#</sup>	A
<b>Escala Si maior</b>	B	C <sup>#</sup>	D <sup>#</sup>	E	F <sup>#</sup>	G <sup>#</sup>	A <sup>#</sup>	B

Figura 27: Escala maior das 7 notas básicas. Fonte: [45].

Em cada escala, cada nota possui uma determinada frequência. Dobrando essa frequência, elevamos essa nota a uma oitava acima, e dividindo pela metade a frequência, rebaixamos uma oitava. Normalmente é adotada cerca de oito ou nove oitavas nas produções musicais padrão. A afinação mais utilizada atualmente é a escala do Lá da 4 oitava, conhecida como Lá de diapasão, e sua frequência tem 440 *Hz*. A escala definida por essa afinação é conhecida como “escala temperada”, onde todas as notas são obtidas a partir do “Lá de diapasão”. Instrumentos com essa afinação são chamados de instrumentos temperados, é o caso, por exemplo, do violão e do piano.

Por opção, a relação entre frequências e notas musicais estabelecida neste trabalho será pela escala temperada. Essa escala pode ser obtida matematicamente por uma progressão geométrica, onde o primeiro termo é a frequência da nota escolhida (em *Hz*) e cuja razão é o valor numérico  $\sqrt[12]{2} \approx 1,0594631$  em decorrência da divisão de uma oitava em 12 intervalos, isto é,  $(1,0594631)^{12} \approx 2$ . Desta forma, obtém-se uma relação entre frequências e notas musicais multiplicando ou dividindo a frequência de uma nota por essa razão. Por exemplo, se o Lá de diapasão tem frequência de 440 *Hz*, então, o Lá<sup>#</sup> nessa oitava terá frequência igual a  $440 \cdot 1,0594631 \approx 466,2$  *Hz*. Portanto, através da razão  $\sqrt[12]{2}$ , é possível construir uma progressão geométrica multiplicando e dividindo sucessivamente esta razão pelo elemento 440, construindo a escala de frequências temperadas em quantas oitavas se queira dentro do espectro sonoro audível. Como ponto de



partida, utiliza-se a Tabela 3 com as frequências associadas as notas musicais em uma escala temperada, nesta tabela, a oitava 0 se inicia em uma frequência de aproximadamente  $16,352 \text{ Hz}$ , que representa a nota  $C_0$ . A velocidade do som usada com referência para obter estas frequências é  $\nu_s = 344 \text{ m/s}$  [45, 46, 9].

Oitava	0	1	2	3	4	5	6	7	8
B	30,87	61,74	123,5	246,9	493,9	987,8	1976	3951	7902
A <sup>#</sup>	29,14	58,27	116,5	233,1	466,2	932,3	1865	3729	7459
A	27,50	55,00	110,0	220,0	440,0	880,0	1760	3520	7040
G <sup>#</sup>	25,96	51,91	103,8	207,7	415,3	830,6	1661	3322	6645
G	24,50	49,00	98,00	196,0	392,0	784,0	1568	3136	6272
F <sup>#</sup>	23,12	46,25	92,50	185,0	370,0	740,0	1480	2960	5920
F	21,83	43,65	87,31	174,6	349,2	698,5	1397	2794	5588
E	20,60	41,20	82,41	164,8	329,6	659,3	1319	2637	5274
D <sup>#</sup>	19,45	38,89	77,78	155,6	311,1	622,3	1245	2489	49,78
D	18,35	36,71	73,42	146,8	293,7	587,3	1175	2349	4699
C <sup>#</sup>	17,32	34,65	69,30	138,6	277,2	554,4	1109	2217	4435
C	16,35	32,70	65,41	130,8	261,6	523,3	1047	2093	4186

Tabela 3: Tabela das frequências temperadas, tendo-se  $A_4 = 440 \text{ Hz}$ . Fonte: [49].

### 3.5.1 Sintetização de uma escala musical temperada

O objetivo nesta seção é criar uma escala temperada a partir da gravação do som de uma nota de um instrumento musical, e capturar no Audacity uma parte da representação gráfica desse som em forma de onda, definida em um período, que caracteriza o timbre desse som. A quantidade de amostras coletadas na parte dessa onda é igual a  $n$ , o que resulta em uma frequência  $f \approx \frac{Fs}{n} \text{ Hz}$ . Para modelar as frequências associadas as notas da escala temperada, como não é prático<sup>21</sup> alterar a taxa de amostragem  $Fs$ , será alterada a quantidade de amostras  $n$  para se obter a frequência que esteja associada a nota musical desejada na escala temperada. A alteração e reconstrução do sinal se dará por uma “Função Spline Cúbica Interpolante” [43].

Dada uma função  $f(x)$  tabelada nos pontos  $x_i, i = 0, 1, 2, \dots, n$ , a função  $S_3(x)$  é chamada spline cúbica interpolante de  $f(x)$  nos nós  $x_i, i = 0, \dots, n$  se existem  $n$  polinômios de grau 3,  $s_k(x), k = 1, \dots, n$  tais que:

- i)  $S_3(x) = s_k(x)$  para  $x \in [x_{k-1}, x_k], k = 1, \dots, n$ ;
- ii)  $S_3(x_i) = f(x_i), i = 0, 1, \dots, n$ ;

<sup>21</sup>Caso se queira exportar os áudios sintetizados digitalmente para um sintetizador de áudio por um cabo midi ou usb, por exemplo, todos os “samples” de áudio devem estar gravados na mesma taxa de amostragem.

- iii)  $s_k(x_k) = s_{k+1}(x_k)$ ,  $k = 1, 2, \dots, (n - 1)$ ;
- iv)  $s'_k(x_k) = s'_{k+1}(x_k)$ ,  $k = 1, 2, \dots, (n - 1)$ ;
- v)  $s''_k(x_k) = s''_{k+1}(x_k)$ ,  $k = 1, 2, \dots, (n - 1)$ .

O primeiro passo é determinar um tempo de áudio  $t$ , em segundos, que todas as notas sintetizadas deverão ter, ou seja, cada áudio que representará cada nota da escala musical deverá ter  $[t \cdot Fs]$  amostras aproximadamente e sendo  $\bar{f}$  a frequência em  $Hz$  da nota musical a ser sintetizada, seu período será replicado  $\bar{f}$  vezes por segundo, portanto, em toda extensão do áudio sintetizado da nota musical, tem-se o período replicado  $t \cdot \bar{f}$  vezes. Na sequência, aplicaremos o algoritmo da Seção 3.3.2 para envelopar o som da nota por uma função exponencial, junto com um complemento para aumentar o volume do áudio. Como a percepção de volume do áudio depende também da frequência, é conveniente que a variação do volume não seja constante e nem linear para todas as notas sintetizadas da escala, pois notas com frequências baixas necessitam de uma ampliação do volume maior que uma nota com frequência alta para se ter a mesma percepção de volume. Pode-se usar como referência para determinar essa variação do volume as curvas isofônicas, disponíveis no gráfico da Figura 10. A modelagem da nota musical do áudio se dará pela seguinte implementação no Scilab.

### 3.5.2 Implementação no Scilab para sintetização de nota musical com envelope exponencial

```

1. clear; stacksize('max'); format(16);
2. m=input("Escreva o nome do arquivo com um período de áudio wav: ","string");
3. f=input("Escreva a frequência da nota musical a ser modelada: ");
4. N=input("Escreva a nota musical a ser modelada: ","string" );
5. b=input("Escreva a amplitude final do áudio em (0,1]:");
6. [W,Fs,bits]=wavread(m);
7. n=strcat([m," em", string(N)]);
8. w=[W(1,:)];
9. t=1:(Fs/f-1)/(length(W)-1):Fs/f;
10. d=splin(t,w,'natural');
11. s=interp(1:Fs/f,t,w,d);
12. z=repmat(s',int(f),1);
13. for i=1:length(z);
14. e(i)=b*exp(i*(log(b/max(z)))/length(z));
15. y(i)=e(i)*z(i);
16. end
17. j=int(100/max(abs(y)));
18. Y=(1+(j-100)/100)*y
19. wavwrite(Y,Fs,n);

```

### 3.5.3 Descrição do algoritmo da implementação no Scilab para síntese de nota musical com envelope exponencial

2. código para pedir um arquivo de áudio com um período de qualquer duração (nesse período estará caracterizado o timbre do áudio);
  3. código para pedir a frequência da nota musical a ser modelada, disponível na Tabela 3.
  4. código para pedir o nome da nota musical a ser modelada, para fazer a concatenação com o nome do arquivo de entrada, para gerar o nome do arquivo de saída.
  5. código para pedir a amplitude final do áudio para o envelopamento exponencial. O valor deve ser maior que 0 e menor ou igual a 1, para simular o som de um instrumento musical, onde, geralmente, em seu envelope dinâmico tem-se a extinção do som com a amplitude instantânea mínima, é conveniente usar um valor próximo de zero.
  6. código para armazenar os dados numéricos do arquivo de entrada na matriz  $W$ .
  9. código para definir uma matriz com  $n$  elementos, um para cada amostra do áudio de entrada, onde o primeiro elemento é 1 e o último é aproximadamente  $\bar{n}$ , igualmente espaçados, obtidos pela fórmula do termo geral de uma progressão aritmética. Para representar a nota a ser sintetizada, com frequência  $\bar{f}$  e na taxa de amostragem  $Fs$ , é necessário  $Fs/\bar{f} \approx \bar{n}$  amostras.
  10. código para obter as derivadas dos dados do arquivo de entrada em função da matriz  $t$  obtida na linha 9.
  11. código para obter as amostras do áudio de entrada no período da nota a ser sintetizada, através da função spline cúbica interpolante com nós nos inteiros entre 1 e  $\bar{n}$ , incluindo os extremos, totalizando  $\bar{n}$  amostras que, na taxa de amostragem  $Fs$ , em  $Hz$ , define-se o período da nota que se deseja sintetizar.
  12. código para replicar as amostras da matriz obtida na linha anterior, fazendo um empilhamento de matriz por coluna, totalizando, aproximadamente,  $[t \cdot Fs]$  amostras.
- Nas linhas 13, 14, 15 e 16 têm-se os códigos para um envelopamento exponencial dos dados obtidos na linha 12, com as réplicas do timbre de entrada no período da nota escolhida.
- Nas linhas 17 e 18 têm-se os códigos para amplificação do volume do áudio envelopado.
19. código para gerar o áudio sintetizado na nota escolhida em formato .wav.

### 3.5.4 Aplicação da implementação no Scilab para síntese de nota musical com envelopamento exponencial

O objetivo nesta seção é criar uma escala temperada a partir da gravação do som da vibração de uma corda do violão na nota Dó da terceira oitava ( $C_3$ ) com afi-

nação na escala temperada, capturando no Audacity uma parte da representação gráfica desse som em forma de onda, definida em um período, caracterizando o timbre desse som (Figura 28), gravado a uma taxa de amostragem de  $44100 \text{ Hz}$ , com  $16 \text{ bits}$  de resolução por amostra, totalizando 337 amostras, de acordo com o valor referente na Tabela 3, que é  $130,8 \approx \frac{44100}{337} \text{ Hz}$ .

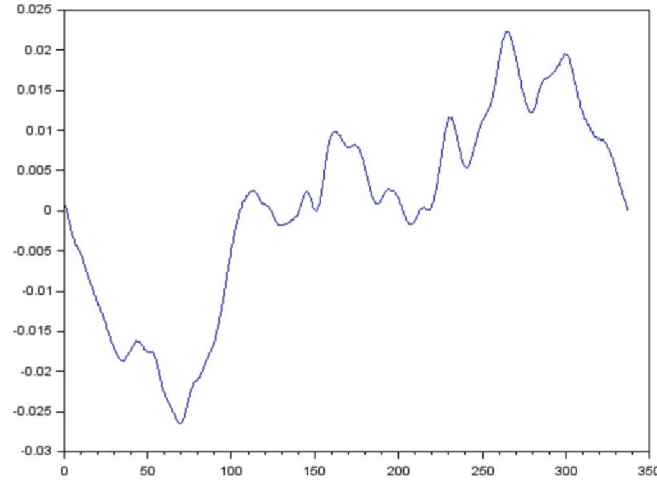


Figura 28: Forma de onda em um período do som de uma corda do violão, na terceira oitava em Dó ( $C_3$ ), na escala temperada. Fonte: captura de tela do Scilab.

Como exemplo de aplicação do algoritmo, vamos sintetizar o som de uma corda de um violão na terceira oitava em Lá ( $A_3$ ). Para isso, é necessário informar a frequência da nota a ser sintetizada, que neste exemplo é  $f = 220 \text{ Hz}$ , valor disponível na Tabela 3. Informa-se também o nome da nota a ser sintetizada para identificar o áudio de saída no formato .wav, que neste exemplo será a nota  $A_3$ . Para simular a extinção do som no envelopamento, vamos definir a amplitude máxima final com o valor  $b = 0,0001$  na escala linear. Após armazenar os dados de entrada, define-se, na linha 9 da implementação, uma matriz com 337 elementos, um para cada amostra do áudio de entrada, onde o primeiro elemento é 1 e o último é aproximadamente 200, igualmente espaçados, obtidos pela fórmula do termo geral de uma progressão aritmética, pois, para representar a nota  $A_3$ , com frequência de  $220 \text{ Hz}$ , na taxa de amostragem de  $44100 \text{ Hz}$ , é necessário  $\frac{44100}{220} \approx 200$  amostras. Na linha 11 vamos obter as amostras do áudio de entrada na frequência da nota  $A_3$ , através da função spline cúbica interpolante com nós nos inteiros entre 1 e 200, incluindo os extremos, totalizando 200 amostras que, na taxa de amostragem de  $44100 \text{ Hz}$ , definirá o período da nota  $A_3$ . Na linha 12 vamos replicar as amostras do vetor obtido na linha anterior, fazendo um empilhamento de matriz por coluna, totalizando  $220 \cdot 200 = 44000 \approx 44100$  amostras, que corresponde ao tempo de áudio de  $\frac{44000}{44100} \approx 1 \text{ s}$ , definido inicialmente



neste exemplo<sup>22</sup>. Prosseguindo com o algoritmo, tem-se um áudio em formato .wav com a sintetização do som da vibração de uma corda do violão na nota  $A_3$ , envelopado por uma função exponencial, cujo gráfico de onda pode ser observado na Figura 29.

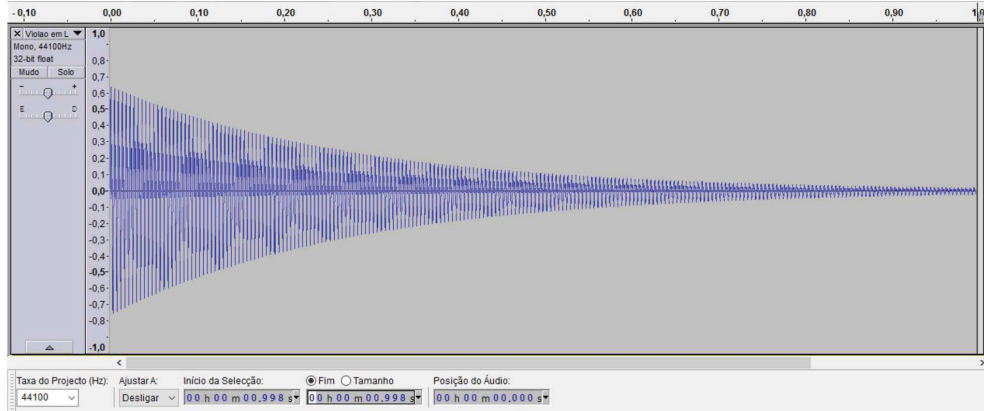


Figura 29: Sintetização do som de um violão na nota  $A_3$  na escala temperada. Fonte: captura de tela do Audacity.

Para obter as demais notas musicais na escala temperada, basta substituir a entrada da frequência da nota  $A_3$  pela frequência de outra nota pretendida, disponível na Tabela 3. Fazendo o mesmo para todas as notas, obtém-se assim a escala temperada sintetizada para o som do violão.

### 3.6 Saída de áudio

Após modelar o som através da manipulação dos dados numéricos do áudio, e gerar um arquivo de áudio em formato digital, precisa-se executar esse arquivo para que o mesmo se transforme em som. Mas, como isso é feito? Como um conjunto de números dispostos em uma matriz se transforma ou volta a ser som?

O que precisa ser feito é o processo inverso da captação do som, através do microfone dinâmico de bobina móvel. Esse processo ocorre pelos alto-falantes, que são formados por componentes que transformam sinais elétricos de uma corrente elétrica em oscilações de pressão no ar, gerando as ondas sonoras. Os componentes básicos de um alto-falante são um ímã permanente, preso na armação do alto-falante, e uma bobina móvel, que está fixa em um cone de papel. Quando uma corrente elétrica variável chega até bobina móvel, esta é repelida ou atraída pelo campo magnético do ímã permanente. Assim, o conjunto bobina-cone é movido para frente e para trás, empurrando o ar em sua volta, criando ondas longitudinais de compressão e rarefação no ar, ou seja, ondas sonoras, e quando

<sup>22</sup>Essa diferença na quantidade de amostras, 100 amostras, se deve ao fato das frequências apresentadas na Tabela 3 ser um valor aproximado.

essas ondas chegam aos nossos ouvidos provocam uma sensação que definimos como sendo o som [47].

## 4 Considerações finais

A fundamentação teórica apresentada neste trabalho tem como objetivo dar uma noção do fenômeno sonoro e da percepção que temos do mesmo, trata-se de uma abordagem superficial, mas suficiente, para compreender os modelos presentes neste trabalho de forma significativa. Sem essa fundamentação, ficaria difícil se libertar da abstração e relacionar os elementos presentes nos modelos com o próprio fenômeno sonoro. Alguns pontos tratados na fundamentação teórica, como no modelo espectral, por exemplo, mesmo sendo pouco aproveitados na modelagem matemática e computacional abordadas, tem como propósito salientar as possibilidades e as alternativas mais viáveis e eficientes para a própria modelagem, e para não criar a impressão de que somente com modelos temporais é que se pode modelar matematicamente as ondas sonoras.

A modelagem matemática e computacional desenvolvida neste trabalho tem suas funcionalidades atestadas com exemplos de aplicações, e possuem as suas vantagens. Por que fazer um modelo para o controle do volume de um áudio, sendo que o mesmo está presente em todos os aparelhos de som? Os modelos que foram desenvolvidos aqui exercem o controle com precisão, tanto do volume, quanto do tempo de reverberação, como também na modelagem de frequências, por exemplo. Uma precisão oferecida pela matemática, que pode ser vista, ouvida e aproveitada para outros fins que envolvam processamento de sinal digital, é a matemática sendo exercida na prática.

A maioria dos educandos possui atualmente um relacionamento muito íntimo com a tecnologia digital, eles manipulam arquivos digitais o tempo todo, escutam músicas e assistem vídeos em formatos digitais, eles compartilham arquivos digitais em uma rede de comunicação por sinal digital, tiram fotos que são armazenadas em arquivos digitais, editam esses arquivos digitais, num tocar de dedos, se divertem com essa tecnologia, como se fosse mágica, e a maioria desses educandos sequer sabe definir o que é um sinal digital. O fascínio que essa tecnologia exerce sobre eles poderia ser explorado para motivá-los a compreendê-la, e quem deveria fazer isso é o professor da disciplina que explica grande parte dessa tecnologia, que é o professor de matemática. Mas o fascínio que se desperta para uns, é o receio que se gera em outros, no receio de se expor, de perder o protagonismo ou por simples comodismo, muitas vezes o professor de matemática evita o tema, e se volta para a obscuridade da abstração matemática, que é o território onde ele se sente mais seguro. Dando a entender que os conteúdos por ele abordados não se relacionam com essa tecnologia, ou que essa tecnologia requer uma matemática muito avançada para ser explorada no nível que se encontra.

Este trabalho é voltado para o professor de matemática do ensino básico, serve

de alicerce para o professor buscar aplicações de vários conceitos matemáticos presentes nesse nível de ensino. Serve para explorar a modelagem matemática em sala de aula aplicada em um fenômeno de interesse comum, abrangendo uma grande variedade de conteúdos da matemática do ensino básico.

O som é um entre tantos outros fenômenos que são pouco ou superficialmente explorados em sala de aula, não por falta de relevância, mas por falta de compreensão do professor das possibilidades inerentes aos conteúdos tradicionalmente trabalhados por ele. Estudar o som somente em ondulatória é um desperdício, foi visto aqui neste trabalho uma infinidade de conteúdos pertinentes ao currículo do ensino básico possíveis de serem relacionados com o fenômeno sonoro, em especial com o áudio digital. A começar pela amostragem, que é o primeiro passo no processo de digitalização de um sinal, sendo um campo significativo para o estudo de sequências. A conversão de sinal analógico para digital e vice-versa é um excelente tema para explorar a relação entre funções contínuas e sequências numéricas, entre outras tantas relações importantes da matemática com o áudio digital e todo universo sonoro, como os decibéis e os logaritmos, as matrizes na representação do sinal discreto, as diversas funções e as relações com os conjuntos numéricos presentes nos modelos aqui trabalhados, os polinômios presentes no processo de reconstrução de um sinal por interpolação, noções de limite e continuidade, conversões de grandezas diversas, até mesmo a geometria espacial da pra se trabalhar usando o som como tema, lembrando que as ondas sonoras são esféricas.

A música tem grande parte de sua teoria com a origem na matemática, a música, ou teoria musical, poderia ser mais explorada na matemática do ensino básico. Por exemplo, nas séries iniciais pode-se explorar a relação entre notas musicais e as frações, e no ensino médio pode-se estabelecer uma ponte entre funções trigonométricas e funções exponenciais, através da relação entre frequências e notas musicais. Essas relações interdisciplinares poderiam estar presentes nos livros didáticos que orientam o professor, e quando estão, aparecem como uma curiosidade, uma nota no final do capítulo, um vislumbre do que pode ser explorado, dando-se mais ênfase a situações problemas sem propósito com as práticas de comum interesse, situações forçadas, fáceis de se entender, mas desinteressantes.

O professor deve ser um agente motivador, deve instigar a curiosidade do aluno, sendo às vezes necessário ir além do conteúdo trabalhado no momento, para trazer algum significado para o aluno. O primeiro passo para aprender é a motivação, não precisa saber tocar um instrumento para gostar de música, mas se a pessoa não gostar de música é pouco provável que ela queira aprender a tocar um instrumento. As pessoas não nascem gostando da matemática, elas aprendem a gostar, e isso acontece quando a mesma traz algum significado para a sua vida.



## Agradecimentos

Agradeço a minha família, em especial a minha esposa e a minha filha, pelo apoio que tive durante o curso. Agradeço aos professores do curso e a instituição de ensino UFPR, pelo conhecimento ofertado. Agradeço também a CAPES pela bolsa de estudo concedida e a SEED por conceder o afastamento com ônus das minhas atividades laborais, sem isso seria inviável a realização deste trabalho.

## Anexo A

### Roteiro para a 1ª parte do pré-processamento: captura do sinal de áudio.

**Software de gravação ou edição de áudio:** Audacity.

**Requisito:** um arquivo de áudio curto, gravado ou não no Audacity, ou um áudio sintetizado gerado pelo próprio Audacity, como, por exemplo, um tom puro (Seção 2.1.7).

**Passo 1:** se já tem o requisito, abri-lo no Audacity. Se não, grave um áudio curto no Audacity definindo os canais de gravação, o tipo de gravação (estereofônica ou monoaural), a taxa de amostragem em  $Hz$ , e a resolução de cada amostra dada em *bits*.

**Passo 2:** obter o conjunto de dados numéricos do arquivo de som aberto no Audacity, temos no Audacity a opção para obter esse conjunto de dados numéricos descritos em um arquivo de texto, em formatos .txt, .csv ou .html, utilizando a seguinte sequência:

- 1º) Escolha na barra de ferramentas do Audacity as opções: **Analisar**  $\Rightarrow$  **Sample data export...** (Figura 30).



Figura 30: Análise de dados de arquivo de áudio no Audacity em formato .wav.  
Fonte: captura de tela do Audacity.

- 2º) Na janela “Sample Data Export...”, Figura 31, temos como definir a forma que se dará a saída dos dados do áudio, nessa janela estão listadas as seguintes opções:



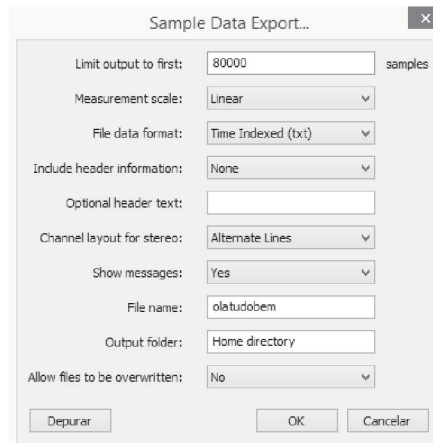


Figura 31: Exportação de dados amostrados, Caixa de diálogo com as opções para exportação de dados amostrados. Fonte: captura de tela do Audacity.

**Limit output to first:** determina quantas amostras serão selecionadas do arquivo de áudio. Para calcular a quantidade total de amostras do arquivo de áudio, deve-se multiplicar o tempo do áudio pela taxa de amostragem utilizada, e preencher a lacuna com um valor igual ou superior a esse produto para abranger toda a extensão do arquivo de áudio.

**Measurement scale:** opções da escala de medição, linear ou dB (decibel).

**File data format:** formato dos dados do arquivo, com as opções:

Sample List (txt): formato de dados .txt com valores de intensidades sonoras dispostos em uma coluna. Indexed List (txt): formato de dados .txt com duas colunas, a primeira com os índices de posicionamento na linha e a segunda coluna com as intensidades sonoras.

Time Indexed (txt): formato de dados .txt com duas colunas, a primeira com os instantes de execução das amostras de intensidades sonoras, e a segunda coluna com as respectivas intensidades sonoras dispostas nas mesmas linhas. Data (csv): formato de dados .cvc, que pode ser lido, por exemplo, em uma planilha eletrônica, com valores de intensidades sonoras dispostos em uma coluna.

Web Page (html): formato de dados .html com descrição do arquivo e dados listados em colunas.

**Include header informations:** opção de redigir cabeçalho.

**Optional header text:** formatos de descrição de dados:

None: somente os dados, sem descrição (necessário para importação do arquivo no Scilab).

Minimal: cabeçalho com taxa de amostragem, com a escala utilizada e posição dos canais.

Standard: cabeçalho com nome do arquivo, quantidade de canais, posição dos canais, taxa de amostragem, tipo de escala utilizada, comprimento da amostra e tempo de execução do áudio.

All: cabeçalho com descrição detalhada.

**Channel layout for stereo:** Disposição de dados para canal estéreo. L -R on Same Line: dados dispostos em duas colunas, a primeira com canal esquerdo, e a segunda com o direito. Alternate Lines: canais esquerdo e direito dispostos em uma coluna com dados alternados. L Channel First: canais esquerdo e direito dispostos em uma coluna, com dados do canal esquerdo primeiro e dados do canal direito na sequência.

**Show messages:** opções de mensagens:

Yes: exibe mensagem com descrição do áudio e caminho do arquivo de saída.

Errors Only: exibir mensagem somente quando houver erros.

None: não exibir mensagem.

**File Name:** nome do arquivo de saída.

**Output Folder:** diretório de saída, pré definido para “home directory”.

**Allow files to be overwritten:** permite ou não que os arquivos sejam substituídos dentro do diretório de destino.

- 3º) Após definir entre as opções como será forma de saída dos dados de áudio, deve-se clicar em “ok”, os dados exportados estarão dispostos em colunas, no formato do arquivo determinado no item anterior (.txt, .csv, ou .html), o arquivo será gerado na pasta do usuário do computador (home directory), com o nome definido no item anterior em (File Name). O arquivo com os dados do áudio exportado poderá ser aberto em um bloco de notas, em uma planilha eletrônica, no GeoGebra, no Scilab, ou em outro aplicativo compatível de interesse.

**Exemplo:** segue um exemplo de exportação de arquivo de áudio .wav <Áudio Estéreo> no Audacity para o formato .txt, com a descrição de áudio da pronúncia “áudio estéreo”. O áudio selecionado para o teste foi gravado em dois canais estéreo, à taxa de amostragem de  $44100\text{ Hz}$ , com  $16\text{ bits}$  de resolução por amostra e tendo  $1,46231\text{ s}$  de gravação, o que totaliza  $[44100 \times 1,46231] = [64487,871] = 64488$  amostras por canal. Como resultado temos quatro colunas de valores, sendo as duas primeiras colunas representando os instantes em segundos e as respectivas intensidades sonoras do primeiro canal, e as duas últimas colunas representando os instantes e as respectivas intensidades do segundo canal. A desvantagem em

fazer esse processo está na imprecisão dos resultados (Figura 32), o que fica evidente pela quantidade de casas decimais utilizadas, que, pela pequena magnitude dos valores obtidos, acarreta em erros consideráveis de aproximação.



Arquivo	Editar	Formatar	Exibir	Ajuda
0.00000	0.00394	0.00000	0.00391	
0.00002	0.00391	0.00002	0.00397	
0.00005	0.00162	0.00005	0.00153	
0.00007	-0.00217	0.00007	-0.00211	
0.00009	-0.00427	0.00009	-0.00436	
0.00011	-0.00388	0.00011	-0.00388	
0.00014	-0.00226	0.00014	-0.00226	
0.00016	-0.00311	0.00016	-0.00314	
0.00018	-0.00839	0.00018	-0.00830	
0.00020	-0.01395	0.00020	-0.01404	
0.00023	-0.01208	0.00023	-0.01208	

Figura 32: Dados do arquivo de áudio exportados para o formato de texto, Áudio com a gravação de voz com a pronúncia “áudio estéreo”. Fonte: captura de tela do bloco de notas.

## Anexo B

### Roteiro para a 2ª parte do pré-processamento: importação dos dados de áudio no Scilab.

Após obter os dados de áudio em arquivo de texto deve-se escolher uma plataforma computacional para proceder a modelagem matemática desses dados, nesse tutorial utilizaremos o Scilab.

#### Modo 1

Importação no Scilab dos dados do áudio em arquivo de texto.

**Software:** Scilab.

**Requisito:** para ser possível importar os dados de áudio no Scilab, o arquivo de texto a ser gerado pelo Audacity deve estar no formato .txt ou .csv, e sem cabeçalho. Na janela principal do Scilab, deve estar selecionado na coluna à esquerda o diretório onde se encontra o arquivo de texto exportado pelo Audacity.

**Exemplo:** como exemplo, vamos usar o arquivo de dados de áudio com o nome “audioestereo” no formato .txt. O código para importar o arquivo no Scilab é  $\langle X = \text{read}('arquivo.txtou.csv', a, b); \rangle$ , onde  $X$  é a matriz que receberá os dados do áudio,  $\langle \text{read} \rangle$  é o código para ler o arquivo, 'arquivo.txt ou .csv' é o nome do arquivo que deve ser escrito junto com a extensão e entre aspas ou apóstrofes. A variável  $a$  indica a quantidade de linhas e  $b$  a quantidade de colunas a serem importadas, para  $\langle a = -1 \rangle$  tem-se todas as linhas importadas.



Figura 33: Importação de dados de um sinal de áudio no Scilab. Console Scilab, código para importação de áudio. Fonte: captura de tela do Scilab.

**Descrição:** Na janela do console do Scilab, (Figura 33), temos como exemplo uma importação de arquivo de dados de sinal de áudio e geração de gráficos das ondas sonoras descritas.

1º) Importação de arquivo de texto.



`<X = read('audioestereo.txt',1,4);>` Código para importar uma matriz com 4 colunas com os dados do áudio.

2º) Exemplos de códigos para separar os dados de áudio: `<t1 = x(:,1);>` Código para criar um vetor t1 dos instantes do áudio da faixa 1 em segundos, a partir da coluna 1 da matriz  $x$ .

`<y1 = x(:,2);>` Código para criar um vetor y1 do sinal das intensidades sonoras do áudio da faixa 1, a partir da coluna 2 da matriz  $x$ .

`<t2 = x(:,3);>` Código para criar um vetor t2 dos instantes do áudio da faixa 2 em segundos, a partir da coluna 3 da matriz  $x$ .

`<y2 = x(:,4);>` Código para criar um vetor y2 do sinal das intensidades sonoras do áudio da faixa 2, a partir da coluna 4 da matriz  $x$ .

`<subplot(211); xtitle('canal 1'); plot(t1,y1); subplot(212); xtitle('canal 2'); plot(t2,y2);>` Código para criar os gráficos de onda das faixas de áudio 1 e 2.

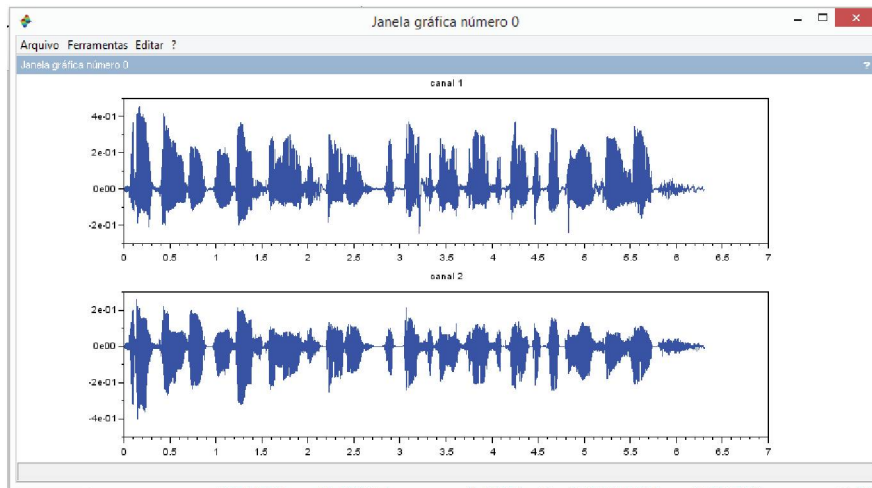


Figura 34: Gráfico de sinal de áudio. Gráfico gerado no Scilab pelo código descrito na Figura 33. fonte: captura de tela do Scilab.

#### Análise inicial dos dados de áudio

A taxa de amostragem, medida em Hertz ( $Hz$ ), é uma das informações mais importantes do conjunto de dados, é obtida dividindo a quantidade de amostras de dados de intensidades sonoras de uma faixa de áudio pelo tempo de execução do áudio. O valor encontrado se refere a quantas amostras foram coletadas em 1 segundo. Por exemplo, para um áudio cujos dados estão representados pelas intensidades sonoras no vetor  $y$  e com um vetor de mesmo tamanho  $t$  para os respectivos instantes de cada intensidade sonora, obtemos a taxa de amostragem  $F_s$  pelo código: `<Fs=int(length(y)/t(length(t)))>`, onde,

`<Int(a)>`: representa a parte inteira do valor de um número  $a$ , visto que, pelas aproximações feitas na coleta de dados, é improvável que a razão seja um número inteiro.

`<length(y)>`: representa o comprimento (tamanho) do vetor  $y$ .

`<t(length(t))>`: código para o último elemento do vetor  $t$ , que indica a duração da gravação do áudio em segundos.

Mesmo não se tratando da forma mais eficiente de obter os dados numéricos de um áudio, esta forma de obtenção dos dados tem suas funcionalidades e é útil, por exemplo, quando se deseja gerar um áudio a partir de uma sequência numérica obtida, por exemplo, em uma planilha eletrônica ou arquivo de texto.

### Modo 3

Importação no Scilab de dados de áudio diretamente do arquivo em formato de áudio digital.

**Requisito:** um arquivo de áudio digital em formato .wav, .mp3 ou .au.

**Software:** Scilab.

**Descrição:** é possível importar arquivos de áudio digital no Scilab de forma mais eficiente que a apresentada no modo anterior. O código de importação no Scilab do arquivo de áudio pode ser feito tanto no console do Scilab como na janela de execução do Scilab. Existem várias alternativas para importar esses dados, por exemplo, para o mesmo arquivo de áudio “Áudio Estéreo” que gerou os dados descritos na Figura 32, segue a descrição de algumas opções:

```
<[Y,Fs,bits]=wavread('audioestereo');>
```

$Y$  é uma matriz de ordem  $i \times j$  que guardará os dados das intensidades sonoras, caso a gravação seja em mais de um canal, os dados de cada canal estarão representados, respectivamente, por cada linha  $i$  da matriz  $Y$  e a quantidade de amostras de cada canal do sinal de áudio estará representada por  $j$ , portanto, todas as faixas de áudio componentes do arquivo de áudio devem ter a mesma quantidade de amostras. A variável  $Fs$  carrega a taxa de amostragem usada durante a gravação e a variável  $bits$  carrega a quantidade de bits para cada amostra definida durante a gravação do áudio no Audacity.

A separação dos canais poderá ser feita através do código `< yk = Y(k,:) >`, sendo  $1 \leq k \leq i$ , isto é,  $y_k$  carrega os dados do canal  $k$ . A importação de dados realizada pelo Scilab pelo código `<wavread>` não carrega o vetor com os instantes de execução das amostras e o tempo de execução do áudio, mas carrega a taxa de amostragem  $Fs$  usada durante a gravação do áudio. Utilizando o código `< t = 1/Fs : 1/Fs : (length(Y(k,:))/Fs); >` é possível obter um vetor  $t$  com os instantes de cada intensidade sonora, sendo a duração do áudio dada pelo código `<t(length(t))>` que busca o instante de execução da última amostra. Podemos

obter a duração do áudio também pelo código  $\langle \text{length}(y(k,:)) * (1/Fs) \rangle$ , para algum  $k$  inteiro, tal que  $1 \leq k \leq i$ . Sendo  $Fs$  a taxa de amostragem, temos que  $\langle 1/Fs \rangle$  representa quantas amostras foram coletadas em um segundo, e sendo as amostras coletadas em um período regular, temos que  $\langle 1/Fs \rangle$  representa o instante da coleta da primeira amostra e também o incremento para obter os instantes das demais amostras.

É possível também importar um arquivo .wav no Scilab através do código  $\langle \text{loadwave} \rangle$ , cujas funcionalidades são parecidas com a do  $\text{wavread}$ . O código básico é:

$\langle X = \text{loadwave}('nomedoarquivo.wav') \rangle$ , sendo que o nome do arquivo deve constar a extensão, e  $X$  representa o vetor ou matriz que armazenará os dados do áudio.

Outra opção para importar um arquivo de áudio é o código  $\langle \text{auread} \rangle$ , com as mesmas funcionalidades do código  $\langle \text{wavread} \rangle$  mas com a vantagem de poder ler outros formatos de arquivos de áudio, como .mp3 e .au, para isso deve-se informar a extensão, exceto para os arquivos com extensão .au. O código básico é:

$\langle [y, Fs, bits] = \text{auread}('nomedoarquivo') \rangle$

Comparando os dois métodos de obtenção de dados de áudio pelo Scilab, constatou-se que, para um mesmo arquivo, e com a mesma taxa de amostragem, os dados obtidos diferem entre si em relação a resolução de cada amostra. Sendo que o primeiro método (exportação do Audacity para arquivo de texto), apresentou como resultado uma baixa e limitada resolução, e sem opção de alteração. Enquanto que pelo segundo método (importação direta pelo Scilab), temos uma melhor resolução de cada amostra. Em ambos os métodos aplicados tivemos um mesmo arquivo primitivo, um áudio no formato “.wav” com  $44100\text{Hz}$  de taxa de amostragem e resolução de  $16\text{bits}$ . Podemos comparar os resultados na captura de tela dada pela Figura 35.

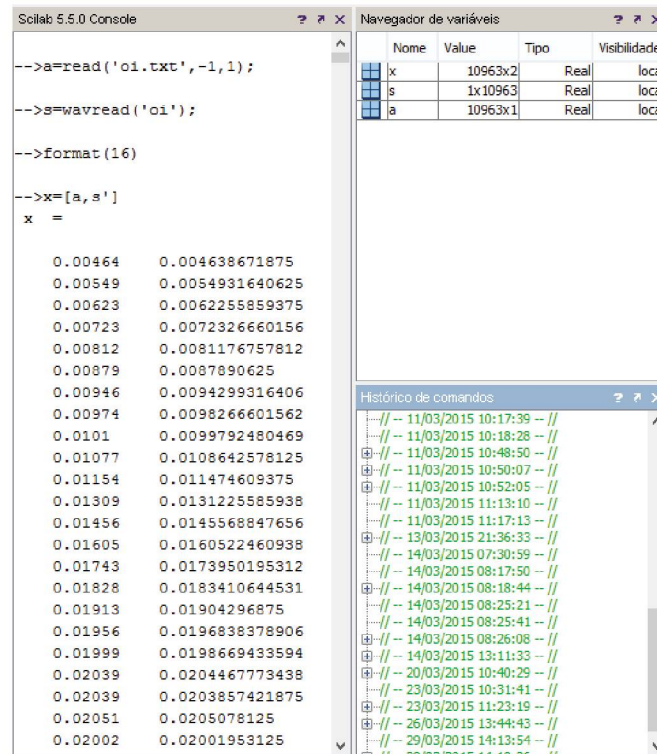


Figura 35: Comparação de dados de áudio via Audacity e Scilab de uma gravação com a descrição “oi”. Fonte: captura de tela no Scilab.

Na Figura 35 temos vetor  $a$  na primeira coluna, com os dados de áudio exportados do Audacity e depois importado pelo Scilab. Observamos também o vetor  $s$  na segunda coluna, com os dados de áudio importados direto do arquivo de áudio .wav para o Scilab. O código `<format(16)>` determina a representação do número com até 16 dígitos, desde que sejam significativos. Pelo resultado, é possível concluir que os dados de áudio importados diretamente pelo Scilab apresentam uma melhor aproximação numérica.

Após manipular os dados de áudio no Scilab temos que converter esses dados em um arquivo de áudio digital para poder ouvir o resultado da manipulação, existem várias formas de gerar um arquivo de áudio no Scilab a partir de um vetor com as intensidades sonoras definidas num intervalo  $[-1, 1]$ . Apresenta-se no Anexo C algumas opções para gerar arquivos de áudio digital a partir de um conjunto de dados numéricos.



## Anexo C

### Roteiro para a geração de um arquivo de áudio no Scilab.

**Requisito:** Um conjunto de dados numéricos no Scilab disposto em matriz, em linha ou coluna, sendo os valores compreendidos no intervalo  $[-1, 1]$ .

**Software:** Scilab

**Descrição:** Depois de efetivada a manipulação do áudio chega a hora de ver os resultados, para isso deve-se gerar um arquivo de áudio a partir dos dados numéricos descritores do áudio, temos no Scilab várias alternativas para isso, segue algumas opções:

**Savewave:** Serve para salvar os dados no formato .wav, o código deve ser da forma:

```
<savewave('nome do arquivo',y,Fs,bits)>
```

O nome do arquivo não deve constar a extensão,  $y$  representa o vetor com intensidades sonoras,  $Fs$  indica a taxa de amostragem a ser adotada (o padrão do Scilab é  $22050Hz$ ), e  $bits$  indica a resolução de cada amostra (o padrão do Scilab é  $16bits$ ).

**Wavwrite:** Tem a mesma funcionalidade do código savewave, o código deve ser da forma:

```
<wavwrite(y,Fs,bits,'nome do arquivo')>
```

**Auwrite:** Tem a mesma funcionalidade dos códigos anteriores, porém, pode gerar arquivos em outros formatos (o formato padrão de áudio é .au), o código é da forma:

```
<auwrite (y, Fs, 'nome do arquivo')>
```

Caso queira outro tipo de extensão, é preciso especificar a extensão no nome do arquivo, por exemplo, “nome do arquivo.mp3”, para arquivo de áudio no formato .mp3.

No Scilab é possível gerar um arquivo de áudio com múltiplos canais, definindo  $y$  na forma de matriz,  $y = [y_1; y_2; \dots; y_n]$ , neste caso, cada conjunto  $y_i$  representa um faixa de áudio descrita por um vetor, sendo as duas primeiras faixas com saída estereofônica, a primeira no canal esquerdo e a segunda no canal direito, e as demais faixas com saída monoaural, isto é, mesmo áudio nos dois canais.

## Referências

- [1] ALONSO, M.; FINN, E. J., Física: um curso universitário, v.2: Campos e Ondas. São Paulo. Editora Edgard Blucher, 1972.
- [2] NUSSENZVEIG, H. M. Curso de Física Básica - v.2, 4. ed. rev., São Paulo. Blucher, 2002.

- [3] RESNICK, R.; HALLIDAY, D.; KRANE, K. S. Física, V.2; 5.ed - Rio de Janeiro: LTC, 2007.
- [4] HANSEN, J. T.; KOEPPEN, B. M. Atlas da fisiologia humana de Netter; Porto Alegre: Artmed, 2003.
- [5] MÖSER, M.; BARROS, J. L. Ingeniería Acústica, Teoría y Aplicaciones. 2.ed. Springer-Verlag Berlin Heidelberg, 2009.
- [6] HALLIDAY, D.; RESNICK, R.; WALKER, J. Fundamentos de Física: gravitação, ondas e termodinâmica, v.2. 8.ed. Editora LTC, 2009.
- [7] LUCINDA, Jair. Reflexões e Ressonâncias, Ondas longitudinais. Disponível em: <<http://reflexoesnoensino.blogspot.com.br/2013/09/ondas-longitudinais.html>>. Acessado em 19 de abril de 2015.
- [8] INFOESCOLA. Ouvido. Disponível em: <<http://www.infoescola.com/wp-content/uploads/2010/08/ouvido.jpg>>. Acessado em 15 de março de 2015.
- [9] MENEZES, Flo. A acústica musical em palavras e sons - Cotia, SP: Ateliê Editorial, 2003.
- [10] UNESP-Acústica e Ruídos -Princípio do som. Disponível em: <<http://www.feb.unesp.br/jcandido/acustica/Apostila/Capitulo%2003.pdf>>. Acessado em 24 de novembro de 2015.
- [11] MATHPAGES. Princípio de Huygens. Disponível em: <<http://www.mathpages.com/home/kmath242/kmath242.htm>>. Acessado em 07 de maio de 2015.
- [12] LOS SENDEROS STUDIO. Onda de pressão sonora. Disponível em: <<http://lossenderosstudio.com/rarefaction.jpg>>. Acessado em 24 de novembro de 2015.
- [13] IAZZETTA, Fernando Henrique de Oliveira. Escola de Comunicação e Artes - USP - Intensidade Sonora. Disponível em: <<http://www2.eca.usp.br/prof/iazzetta/tutor/acustica/intensidade/intensidade.htm>>. Acessado em 07 de maio de 2015.
- [14] IAZZETTA, Fernando Henrique de Oliveira. Escola de Comunicação e Artes - USP - Pressão Sonora. Disponível em: <<http://www2.eca.usp.br/prof/iazzetta/tutor/acustica/intensidade/pressao.htm>>. Acessado em 07 de maio de 2015.
- [15] IAZZETTA, Fernando Henrique de Oliveira. Escola de Comunicação e Artes - USP - Potência Sonora. Disponível em: <<http://www2.eca.usp.br/prof/iazzetta/tutor/acustica/intensidade/potencia.htm>>. Acessado em 07 de maio de 2015.

- [16] SOM AO VIVO. O decibel e seus mistérios. Disponível em: <<http://www.somaovivo.org/artigos/o-decibel-e-seus-misterios-parte-i/>>. Acessado em, 10 de maio de 2015.
- [17] CONDE, Flavio - Redes de microfones em tempo real: Uma implementação com hardware de baixo custo e software de código aberto. ed.rev. São Paulo, 2010.
- [18] SPADA, Adriano Luiz. Microfones - Artigos técnicos. Disponível em: <[http://www.attack.com.br/artigos\\_tecnicos/mic\\_1.pdf](http://www.attack.com.br/artigos_tecnicos/mic_1.pdf)>. Acessado em 15 de maio de 2015.
- [19] ZURMELY, Roland M. Espectro. Disponível em: <<http://www.qsl.net/py4zbz/teoria/espectro.htm>>. Acessado em 12 de outubro de 2015.
- [20] ZURMELY, Roland M. Digitalização de um sinal analógico. Disponível em: <<http://www.qsl.net/py4zbz/teoria/digitaliz.htm>>. Acessado em 12 de outubro de 2015.
- [21] DAMPER, R. I. - Introduction to Discrete-Time Signals and Systems, Department of Electronics and Computer Science University of Southampton - UK, Chapman and Hall, 1995.
- [22] DEE, UFRN. PCM - Pulse Code Modulation - Sistemas de Telecomunicações - Departamento de Engenharia Elétrica - UFRN. Disponível em: <<http://www.dee.ufrn.br/pcm.pdf>>. Acessado em 06 de maio de 2015.
- [23] LERNER, Sandy. The dilettante's dictionary - Quantization. Disponível em: <<http://www.dilettantesdictionary.org/defimages/quantization.gif>>. Acessado em 14 de junho de 2015.
- [24] KREYSZIG, E. - Matemática Superior para Engenharia - vol.2, 9.ed., RJ, LTC, 2009.
- [25] KUMAR, Subodh. Aliasing. Disponível em: <<http://www.cs.jhu.edu/subodh/457/images/alias.gif>>. Acessado em 12 de outubro de 2015.
- [26] FIGUEIREDO, Djairo Guedes. Análise de Fourier e Equações Diferenciais Parciais. Rio de Janeiro, Instituto de Matemática Pura e Aplicada, CNPq, 1977.
- [27] NEVES, Armando G. M. Transformada de Fourier, DMAT, UFMG. Disponível em: <<http://www.mat.ufmg.br/aneves/ensino/edb/chapter8.pdf>>. Acessado em 18 de outubro 2015.

- [28] CARVALHO, Paulo Cezar P. - Métodos Matemáticos e Computacionais em Música - São Carlos, SP: SBMAC, 2009.
- [29] CASTRO, Maria Cristina Felippetto De. Teorema de Nyquist - Teorema da Amostragem. Disponível em: <[http://www.feng.pucrs.br/decastro/TPI/TPI\\_Cap3\\_parte2.pdf](http://www.feng.pucrs.br/decastro/TPI/TPI_Cap3_parte2.pdf)>. Acessado em 17 de novembro 2015.
- [30] LA WEB DE FÍSICA-Princípio de Huygens. Disponível em: <<http://forum.lawebdefisica.com/entries/439-Principio-de-propagacao-de-Huygens-Fresnel>>. Acessado em 22 de novembro 2015.
- [31] RAENG. The Royal Academy of Engineering - The Study of Root Mean Square (RMS) Value - Disponível em: <<http://www.raeng.org.uk/publications/other/8-rms>>. Acessado em 22 de novembro de 2015.
- [32] MEC. Banco Internacional de Objetos Educacionais - Interferência em Ondas Transversais e Longitudinais. Disponível em: <<http://objetoseducacionais2.mec.gov.br/bitstream/handle/mec/23266/Interferencia%20em%20Ondas%20Transversais%20e%20Longitudinais.jpg?sequence=1>>. Acessado em 22 de novembro de 2015.
- [33] NAS. National Academy of Sciences of the United States of America - Biographical Memoir of Alexander Graham Bell - by Harold S. Osborne. Disponível em: <<http://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/bell-alexander-graham.pdf>>. Acessado em 22 de novembro de 2015.
- [34] Donoso, José Pedro - História da Acústica 1 - Acústica Física - Instituto de Física de São Carlos - USP. Disponível em: <<http://www.gradadm.ifsc.usp.br/dados/20142/FFI0210-1/Historia%20da%20Acustica%201.pdf>>. Acessado em 22 de novembro de 2015.
- [35] FOURIER, Joseph, translated by Alexander Freeman (published 1822, translated 1878, re-released 2003). The Analytical Theory of Heat. Dover Publications. 2003 unabridged republication of the 1878 English translation by Alexander Freeman of Fourier's work *Théorie Analytique de la Chaleur*, originally published in 1822.
- [36] EAC, UfSM. Qualidade Sonora - Pesquisa - Engenharia Acústica - UFSM - Disponível em: <<http://www.eac.ufsm.br/pesquisa/qualidade-sonora>>. Acessado em 22 de novembro de 2015.



- [37] RUMSEY, Francis; RUMSEY, McCormick Tim. Sound and recording. 6<sup>th</sup> ed., 1954. Disponível em: <[http://www.adnicholson.com /Sound%20and%20Recording,%206th%20Edition.pdf](http://www.adnicholson.com/Sound%20and%20Recording,%206th%20Edition.pdf)>. Acessado em 22 de novembro de 2015.
- [38] ADAMS, Mike. Lee de Forest - King of Radio, Television, and Film - San Jose State University. Disponível em: <[https://www.smpte.org/sites/default/files/25-1600-TS9-1-ST03-FINAL.Adams\\_.pdf](https://www.smpte.org/sites/default/files/25-1600-TS9-1-ST03-FINAL.Adams_.pdf)>. Acessado em 22 de novembro de 2015.
- [39] HIGUTI, Ricardo; KITANO, Cláudio - Sinais e Sistemas - Princípios de Comunicações - DEE-FEIS-UNESP, 2003. Disponível em: <[http://www.feis.unesp.br/Home/departamentos/engenhariaeletrica /op-toelettronica/sinais\\_e\\_sistemas.pdf](http://www.feis.unesp.br/Home/departamentos/engenhariaeletrica/op-toelettronica/sinais_e_sistemas.pdf)>. Acessado em 22 de novembro de 2015.
- [40] HELMOHOLTZ, Hermann. On the Sensations of Tone (Die Lehre con den Tonempfindungen - 1862). New York, Dover Publications, 1954.
- [41] MIDI, The Association. Envelope dinâmico - Disponível em: <[http://www.midi.org/images/tut\\_fig7.gif](http://www.midi.org/images/tut_fig7.gif)>. Acessado em 23 de novembro de 2015.
- [42] FRADEN, Jacob. Handbook of Modern Sensors. 4.ed. Nova Iorque: Editora Springer Science and Business Media, 2010.
- [43] RUGGIERO, M. A. G., LOPES, V. L. da R. Cálculo Numérico: Aspectos Teóricos e Computacionais. Editora Pearson Education. 1996.
- [44] WILLIAMS, Rich L. Practical music production. Disponível em: <<http://www.practical-music-production.com/images/reverb.jpg>>. Acessado em 12 de abril de 2015.
- [45] DESCOMPLICANDO A MÚSICA. Escalas musicais. <<http://www.descomplicandoamusica.com/escalas-musicais/>>. Acessado em 18 de janeiro de 2016.
- [46] IAZZETTA, Fernando Henrique de Oliveira. ECA - USP - Tabela de Frequências, Períodos e Comprimentos de Onda. Disponível em: <<http://www2.eca.usp.br/prof/iazchetta/tutor/acustica/introducao/tabela1.html>>. Acessado em 18 de janeiro de 2016.
- [47] SILVA, Domiciano Corrêa Marques da - Alto-falantes. Disponível em: <<http://mundoeducacao.bol.uol.com.br/fisica/altofalantes.htm>>. Acessado em 20 de janeiro de 2016.
- [48] SCILAB, Manual. Disponível em: <[https://help.scilab.org/docs/5.3.0/pt\\_BR/](https://help.scilab.org/docs/5.3.0/pt_BR/)>. Acessado em 12 de abril de 2015.

- [49] CAMPBELL, Murray; GREATER, Clive. The Musician's Guide to Acoustics. New York, Oxford University Press, 2001. In: Menezes, Flo - A acústica musical em palavras e sons-p.268. Cotia, SP: Ateliê Editorial, 2003.